UNIVERSITÉ DE MONTRÉAL

CENTRE DE RECHERCHE SUR LES TRANSPORTS

MÉMOIRE D'ÉTUDIANT

CASE POSTALE 6128, STATION A, MONTRÉAL, QUÉBEC, H3C 3J7

TRANSFORMATIONS DE BOX-COX ET DEMANDE DE TRANSPORT EN COMMUN A MONTREAL

par
Lisa Di Pietro

Ce travail a été subventionné par le Centre de recherche et de développement de Transports Canada et par le programme F.C.A.C. du Ministère de l'éducation du Québec

Université de Montréal
Centre de recherche sur les transports - Publication # 155
Août 1979

Dépôt légal - 1er trimestre 1980

RESUME

L'objet du présent rapport de recherche est de montrer d'une manière empirique que les transformations monotoniques Box-Cox de puissance directe et inverse peuvent être utiles pour déterminer la spécification stochastique d'un modèle et plus précisément, pour caractériser la forme fonctionnelle de l'hétéroscédasticité des résidus. La formulation générale de l'hétéroscédasticité proposée par les Professeurs Gaudry et Dagenais est insérée dans un modèle linéaire et dans un modèle où les variables endogène et exogènes sont également sujettes aux transformations de puissance directe. De plus, la détermination simultanée de la forme de l'hétéroscédasticité et de la forme fonctionnelle de la partie systématique est incorporée à une structure d'autocorrélation multiple des résidus.

Dans cette étude, le cadre analytique choisi est assujetti à une fonction de demande de transport en commun à Montréal. Les déplacements des voyageurs (plus précisément des écoliers) sont expliqués à partir de données en séries chronologiques pour les années 1956 à 1971.

Nos résultats nous permettent, à partir des divers instruments utilisés, d'établir une distinction explicite entre la non-linéarité et l'hétéroscédasticité du modèle. De plus, ils montrent que la détermination conjointe des parties systématique et stochastique du modèle peut avoir des effets considérables sur les estimations des paramètres du système et s'avère supérieure à des procédures moins générales.

UNIVERSITÉ DE MONTRÉAL

TABLE DES MATIERES

RESUME i

INTRODUCTION 1

CHAPITRE I : Revue des principales études traitant de la forme fonctionnelle de la partie fixe et de la spécification du terme stochastique 3

1.1. Schéma classique 4

1.1.1. Analyse systématique de la partie aléatoire du modèle 4

A) Analyse de l'hétéroscédasticité 5
B) Analyse de l'autocorrélation 7

1.2. Etude systématique de la formulation générale de la partie fixe du modèle 8

. Transformation Box-Cox 9

1.3. Etude systématique de la formulation générale de la partie fixe et de la partie aléatoire du modèle 10

1.3.1. Formulation de l'hétéroscédasticité dans un modèle linéaire 13

1.3.1.1. Une forme générale pour les erreurs hétéroscédastiques 13

1.3.1.2. Fonction de vraisemblance pour un modèle linéaire 16

1.3.1.3. Fonction de vraisemblance pour un modèle non-linéaire 17

1.3.2. Formulation de l'hétéroscédasticité et de l'autocorrélation du premier ordre dans un modèle non-linéaire 18

. Fonction de vraisemblance en présence de l'hétéroscédasticité et de l'autocorrélation du premier ordre 19

1.3.2.1. Formulation simplifiée en présence de l'autocorrélation du premier ordre 21

1.3.2.2. Formulation simplifiée en présence de l'autocorrélation multiple 22

CHAPITRE II : Procédure d'estimation 25

CHAPITRE III : Modèle à estimer 32

3.1. Application à un modèle de demande de transport en commun . . . 32

3.1.1. Choix de la paramétrisation de la partie systématique du modèle 34

3.1.2. Choix de la paramétrisation de la partie stochastique du modèle 37

3.2. Nomenclature des hypothèses 37

3.2.1. Hypothèses sur la partie systématique du modèle 38

3.2.2. Hypothèses sur la partie stochastique du modèle 39

3.3. Comparaison des hypothèses 43

. en absence d'autocorrélation 43
. en présence d'autocorrélation 44

CHAPITRE IV : Analyse des résultats 49

4.1. Justification de l'opérationnalité du modèle 49

4.2. Estimation et tests de la forme fonctionnelle 52

4.2.1. Justification pratique de l'utilité de la détermination conjointe par une comparaison des modèles à partir des hypothèses émises sur la partie aléatoire . . 52

4.2.2. Justification de l'utilité de la généralisation de la paramétrisation de la forme fonctionnelle par une comparaison des modèles à partir des hypothèses émises sur la partie systématique 63

4.2.3. Détermination du modèle optimal pour la fonction de demande du transport en commun à Montréal (marché des écoliers) 69

4.3. Divers effets de la forme fonctionnelle de la partie systématique et de la spécification de l'erreur 70

4.3.1. Effets de la forme fonctionnelle et de la spécification de l'erreur sur la distribution des erreurs résiduelles autorégressives 70

. Coefficient de détermination multiple 71

4.3.2. Effets de la forme fonctionnelle de la partie systématique et de la spécification de l'erreur sur la signification des variables 73

4.3.3. Effets de la forme fonctionnelle de la partie systématique et de la spécification de l'erreur sur les élasticités 83

CONCLUSION 92

APPENDICE A : Logarithme de la fonction de vraisemblance concentrée . . 94

APPENDICE B : Hypothèse de normalité du terme stochastique 98

APPENDICE C : Mesures du coefficient de corrélation multiple 101

APPENDICE D : Définition des variables 103

BIBLIOGRAPHIE 106

LISTE DES TABLEAUX

TABLEAU 1 : Nombre d'observations positives pour les différentes variables 36

TABLEAU 2 : Tableau récapitulatif des diverses hypothèses émises . . 41

TABLEAU 3 : Nombre d'itérations, temps d'exécution et coûts monétaires 52

TABLEAU 4 : Valeurs de λ_y, λ_{xk}, ρ_ℓ, λ_{em} et δ_m pour $X_m \equiv$ *SHONF* et valeurs correspondantes des fonctions de vraisemblance. Séries * et A 54

TABLEAU 5 : Valeurs de λ_y, λ_{xk}, ρ_ℓ, λ_{em} et δ_m pour $X_m \equiv$ *TEND* et valeurs correspondantes des fonctions de vraisemblance. Séries * et A 57

TABLEAU 6 : Valeurs de λ_y, λ_{xk}, ρ_ℓ, λ_{em} et δ_m pour $X_m \equiv$ *TEND* et valeurs correspondantes des fonctions de vraisemblance. Séries * et A 60

TABLEAU 7 : Valeurs de λ_y, λ_{xk}, ρ_ℓ, λ_{em} et δ_m pour $X_m \equiv$ *SHONF* et *TEND* et valeurs correspondantes des fonctions de vraisemblance. Séries A et AA 62

TABLEAU 8 : Valeurs de λ_y, λ_{xk}, ρ_ℓ, λ_{em}, δ_m et valeurs correspondantes des fonctions de vraisemblance. - Modèle BC_SAU - Séries *, A, B, C, D, E 64

TABLEAU 9 : Valeurs de λ_y, λ_{xk}, ρ_ℓ, λ_{em}, δ_m et valeurs correspondantes des fonctions de vraisemblance. - Modèle BC_SHG_2 - Séries *, A, B, C, D, E 65

TABLEAU 10 : Valeurs des tests statistiques χ^2 pour les paramètres λ_y et λ_{xk}. - Modèle BC_SAU - Séries *, A, B, C, D, E . . . 66

TABLEAU 11 : Valeurs des tests statistiques χ^2 pour les paramètres λ_y et λ_{xk}. - Modèle BC_SHG_2 - Séries *, A, B, C, D, E . . 66

TABLEAU 12 : Valeurs des paramètres qui optimisent la fonction de vraisemblance 69

TABLEAU 13 : Fonctions d'autocorrélation et d'autocorrélation partielle en différence 0 des résidus 72A

TABLEAU 14 : Valeurs des coefficients paramétriques, des élasticités, des tests statistiques, de L^{***} pour divers modèles 74A

TABLEAU 15 : Nombre de paires de corrélation partielle évaluées à différents niveaux pour certains modèles 82

TABLEAU 16 : Rapport des élasticités du temps d'attente sur le temps de transport selon divers modèles 88

TABLEAU 17 : Valeur implicite des prix du temps ($/heure) 89

INTRODUCTION

L'économétrie, du moins jusqu'à tout récemment, a étudié et estimé des formes structurelles pour lesquelles la spécification du terme aléatoire est déterminée d'une manière soit rationnelle, soit irrationnelle. Très souvent et à tort, des modèles de demande (ou d'offre) à équation(s) simple ou simultanées ou encore des systèmes complets de demande sont estimés en ne respectant pas la vraie nature de la variable stochastique. De même, la détermination de la forme fonctionnelle de la partie systématique de ces modèles ou systèmes est souvent modulée sur une base arbitraire ou encore appuyée sur des études empiriques antérieures qui ne justifient pas leur utilisation. L'introduction des transformations paramétriques a cependant résorbé, d'une certaine façon, les problèmes inféodés à une mauvaise spécification de la forme fonctionnelle. Nonobstant, la paramétrisation de la partie fixe ne constitue pas toujours en elle-même un instrument suffisant pour assurer la neutralisation de l'hétéroscédasticité des erreurs résiduelles. Pour réaliser cet objectif, nous devons incorporer dans la fonction considérée l'instrument qui permet d'infléchir dans le sens de la respectabilité, une des hypothèses fondamentales inhérentes à la variable aléatoire à savoir l'homoscédasticité des résidus. La théorie classique suggère plusieurs formulations (celles-ci sont exposées dans le premier chapitre) arbitraires de correction pour l'hétéroscédasticité. Cependant, comme ces formulations sont déterminées d'une manière contraignante, elles peuvent constituer des outils de correction insuffisants. La nécessité d'introduire une formulation générale de l'hétéroscédasticité, c'est-à-dire une formulation qui incorpore la procédure Box-Cox à la variable aléatoire s'avère donc primordiale. La minimisation de l'erreur ou encore la maximisation de la vraisemblance requiert un modèle qui incorpore le moins de restrictions possibles et par conséquent la procédure Box-Cox doit

être généralisée aux erreurs qui suivent un processus autorégressif d'ordre multiple si les données utilisées sont constituées à partir de séries chronologiques.

Le présent travail a pour objet l'étude d'une formulation générale de l'hétéroscédasticité; celle-ci, proposée par Gaudry et Dagenais (1978), est insérée dans un modèle où les transformations Box-Cox sont également appliquées à la partie fixe du modèle. La détermination simultanée de la forme fonctionnelle des parties fixe et aléatoire, son estimation par la procédure du maximum de vraisemblance, son utilité, ses effets sur la distribution des résidus, sur la signification des variables et sur les élasticités sont considérés et analysés d'une façon exhaustive.

CHAPITRE I

REVUE DES PRINCIPALES ETUDES TRAITANT DE LA FORME FONCTIONNELLE DE LA PARTIE FIXE ET DE LA SPECIFICATION DU TERME STOCHASTIQUE.

L'introduction systématique des considérations de la forme fonctionnelle flexible dans les modèles de régression est relativement récente dans la littérature.

L'application de la transformation paramétrique Box-Cox sur les variables de la partie fixe du modèle a pour essence la minimisation de l'erreur, laquelle erreur reflète une partie de la mauvaise spécification de la forme fonctionnelle. L'imposition de restrictions établies à priori sur les paramètres monotoniques, v.g. une forme linéaire ou logarithmique, amène des propriétés de comportement (élasticité constante dans une forme double-log) qui ne sont pas sujettes à une justification. Le processus d'estimation va générer la forme fonctionnelle, ce qui confère à la transformation paramétrique un attribut très puissant. La transformation de puissance n'a par aucun moyen la prétention de trouver le "vrai" modèle; cependant, elle contribue à déterminer la forme fonctionnelle qui donne le "meilleur" ajustement.

Dans la littérature courante, la méthode d'estimation utilisée est celle du "scanning", de sorte qu'une famille paramétrique de fonctions est établie. Dans ce cas, tout comme dans le cas où une méthode de maximisation non-linéaire est utilisée, nous n'avons pas à préciser la forme fonctionnelle "exacte" à priori.

La détermination des hypothèses sur la partie aléatoire du modèle telles que développées dans la plupart des écrits est également sujette à discussion. Dans la littérature classique, la spécification de la variable

aléatoire est explicitement fondée sur des postulats qui ne justifient pas nécessairement leur utilisation. Une utilisation rationnelle du terme de l'erreur exige une détermination beaucoup plus générale que celle développée dans la théorie classique. La généralisation de la procédure Box-Cox à la partie aléatoire va permettre de caractériser une forme fonctionnelle de l'hétéroscédasticité des résidus plus générale.

Nous allons dans la première section de ce chapitre présenter les postulats fondamentaux de la théorie classique traditionnelle après quoi, nous allons étudier l'application de la transformation Box-Cox aux parties fixe et stochastique du modèle de régression.

1.1. Schéma classique.

1.1.1. Analyse systématique de la partie aléatoire du modèle.

Considérons la relation linéaire générale suivante:

$$y_t = \sum_k \beta_k X_{k_t} + u_t \quad , \quad \text{pour } t=1,\ldots,n, \text{ où } n \text{ est le nombre d'observations.} \tag{1}$$

Les n équations peuvent être exprimées dans la notation matricielle suivante:

$$y = X\beta + u \ , \tag{2}$$

où y est un vecteur $(n \times 1)$ d'observations sur la variable endogène,

X est une matrice $(n \times k)$ non stochastique de rang k,

β est un vecteur $(k \times 1)$ de paramètres. (La constante β_o est ignorée dans notre modèle afin de faciliter les buts expositoires.)

u est un vecteur $(n \times 1)$ d'erreurs aléatoires distribuées normalement et identiquement avec une espérance mathématique nulle, $E(u)=0$ et une matrice variance-covariance, $E(uu')=\psi^2 I$ (où I est une matrice d'identité d'ordre $n \times n$).

Cette matrice variance-covariance laisse supposer

a) que la variable aléatoire est homoscédastique, c'est-à-dire que la variance de l'erreur est constante et finie

$$E(u_t^2) = \psi^2 \quad , \quad \text{pour tout } t$$

(Ce postulat est nécessaire (mais non suffisant) si nous voulons obtenir par la méthode des moindres carrés ordinaires, des estimateurs à variance minimale.)

et,

b) que les erreurs ne sont pas autocorrelées entre elles, c'est-à-dire

$$E(u_t u_{t+s}) = 0 \quad , \quad \text{pour } s \neq 0.$$

Chacune de ces hypothèses doit être respectée si nous désirons estimer l'équation (2) par les moindres carrés ordinaires[1].

Si l'un de ces postulats fondamentaux n'est pas respecté, la matrice de variance-covariance ne s'identifiera plus comme étant égale à $\psi^2 I$, mais elle deviendra égale à $\psi^2 V$, $E(uu')=\psi^2 V$ (où V est une matrice symétrique définie positive d'ordre $n \times n$) $= \Omega$.

Les hypothèses sur la forme de Ω nous permettent de distinguer entre les éléments diagonaux et les éléments non-diagonaux de la matrice.

A) Analyse de l'hétéroscédasticité.

Si nous postulons la présence de l'hétéroscédasticité dans les résidus et l'absence d'autocorrélation entre les erreurs résiduelles, la matrice Ω prendra la forme suivante:

1.- L'estimation par les moindres carrés ordinaires lorsque les propriétés (a) et (b) ne sont pas respectées génère †) des paramètres inefficaces, quoique non-biaisés,
††) des écarts-types biaisés.

$$E(u_t u_{t'}) = \psi^2 V = \psi^2[f(X_{1_t},\ldots,X_{M_t})] \quad , \quad \text{pour } t=t' \qquad (3)$$

où ψ^2 est inconnu et où $f(X_{1_t},\ldots,X_{M_t})$ est une fonction de plusieurs variables exogènes,

$$= 0 \quad , \quad \text{pour } t\neq t'. \qquad (3')$$

En général, dans la littérature, l'analyse de l'hétéroscédasticité est concentrée sur des variantes particulières de la matrice diagonale. Très souvent, la fonction mathématique qui compose les éléments diagonaux de la matrice V emprunte une forme où la variance du terme d'erreur est proportionnelle au carré de la variable indépendante X.

Cette formulation d'hétéroscédasticité constituée d'une seule variable est caractérisée par le concept d'égalité suivant:

$$E(u_t^2) = \psi^2 X_{k_t}^2. \qquad (4)$$

La présence[2] de la variable explicative peut être identifiée par le test paramétrique de Goldfeld et Quandt (1965, pp.539-547).

Parmi les autres formulations d'hétéroscédasticité élaborées dans les différents ouvrages, nous pouvons spécifier celles appartenant à des auteurs tels que Theil (1951, cité dans Goldfeld et Quandt, 1972, p.82), Glejser (1969, pp.316-323) et autres.

L'article de Theil est basé sur l'hypothèse que l'hétéroscédasticité prend la forme

---

2.- Outre le test de Goldfeld-Quandt, de nombreux autres tests non-constructifs (ces tests de détection servent à établir la présence ou l'absence de l'hétéroscédasticité sans tenir compte de l'estimation) ont été mis en plan. Parmi ces tests, nous pouvons distinguer ceux de Ramsey, Pearson, Barlett et autres (cités dans Goldfeld et Quandt, 1972, pp.85-91). En plus de cette approche paramétrique, la littérature suggère deux autres approches: l'approche basée sur le test non-paramétrique "peak-test" de Goldfeld-Quandt et l'approche heuristique qui inclut les graphiques des résidus.

$$E(u_t^2) = \gamma[E(y_t)]^2, \tag{5}$$

c'est-à-dire que la variance de l'erreur est proportionnelle au carré de la moyenne de la variable dépendante, y.

En ce qui a trait à Glejser, il a proposé diverses formes qui se limitent à estimer (par tâtonnements) l'hétéroscédasticité dans le cas d'une variable explicative unique en supposant une fonction explicite. Les formes retenues et pour lesquelles nous ne connaissons pas à priori la spécification correcte de l'erreur sont apparentées à:

$$\begin{aligned} |u_t| &= a_0 + b_0 X_{k_t} \\ |u_t| &= \sqrt{a_0 + b_0 X_{k_t}} \\ |u_t| &= a_0 + b_0 X_{k_t}^2 \end{aligned} \tag{6}$$

B) Analyse de l'autocorrélation.

Si nous postulons que les erreurs sont homoscédastiques et qu'elles sont correlées entre elles, la matrice Ω s'identifiera par

$$E(u_t u_{t'}) = \psi^2 \quad , \quad \text{pour } t=t' \tag{7}$$

$$\neq 0 \quad , \quad \text{pour } t\neq t' \tag{7'}$$

où u_t est généré par un processus autorégressif spécifié comme étant:

$$u_t = \sum_{\ell=1}^{r} \rho_\ell \, u_{t-\ell} + e_t \quad , \quad \text{pour } t=1,\ldots,n. \tag{8}$$

Nous supposons que les variables aléatoires e_t répondent aux hypothèses de normalité classiques habituelles, c'est-à-dire qu'elles sont distribuées identiquement et indépendamnent avec une moyenne zéro, $E(e_t)=0$ et une matrice variance-covariance, $E(e_t e_{t'})=\sigma^{**2} I$ et que le schéma autorégressif satisfait l'hypothèse stationnaire standard, $|\Sigma\rho| < 1$.

Dans la pratique habituelle la spécification de la structure autoré-

gressive, tout comme les formes d'hétéroscédasticité, est souvent modifiée indépendamment de la forme fonctionnelle de la partie fixe.

1.2. Etude systématique de la formulation générale de la partie fixe du modèle.

Dans la théorie économique conventionnelle, il existe une certaine ambiguïté lorsqu'il s'agit de choisir d'une manière rationnelle la forme mathématique appropriée de la partie fixe. De façon générale, les fonctions économiques proposées par les différents auteurs et inhérentes à des théories diverses (théorie monétaire et autres) sont estimées à partir de formes fonctionnelles spécifiques déterminées arbitrairement ou encore,appuyées sur des études empiriques antérieures. Les formes restrictives les plus communes se limitent presque essentiellement aux formes linéaire, logarithmique[3] et semi-logarithmique.

Si aucune théorie spécifie à priori le degré d'optimalité d'une fonction ou si une telle théorie existe et devrait être vérifiée empiriquement, il est non utopique d'espérer atteindre ce degré d'optimalité en utilisant à cette fin des transformations de puissance paramétriques. En effet, cette alternative semble très séduisante pour optimiser la forme fonctionnelle puisqu'elle permet d'exprimer le problème d'estimation dans une forme aussi générale que possible et qu'elle laisse les données imposer elles-mêmes les restrictions sur cette forme.

La prise en considération des transformations monotoniques sur les variables positives a été étudiée successivement par Anscombe et Tukey (1954,

---

3.- La spécification de la forme fonctionnelle logarithmique est souvent utilisée dans la littérature parce qu'elle permet d'anticiper l'homoscédasticité des erreurs résiduelles. Cependant, il n'est pas assuré que la forme fonctionnelle appropriée pour le modèle soit vraiment log-linéaire (Gaudry et Dagenais, 1978, p.1).

cité dans Gaudry et Dagenais, 1978, p.3), Tukey (1957), Turner et autres (1961), Box et Tidwell (1962) et Box et Cox (1964).

. Transformation Box-Cox.

Si la variable dépendante et les variables indépendantes de l'équation (1) étaient affectées par des paramètres de puissance λ, l'équation ainsi transformée aurait la forme structurelle suivante:

$$y_t^{(\lambda_y)} = \sum_k \beta_k X_{k_t}^{(\lambda_{xk})} + u_t \qquad (9)$$

où λ_y représente un scalaire,

λ_{xk} représente un élément du vecteur λ_x,

β_k est un élément du vecteur β.

La transformation Box-Cox[4] de chaque variable Z_t, si Z_t symbolise la variable dépendante y_t ou la k ième variable indépendante "non Bolean" X_{k_t}, est définie par le paramètre λ comme étant

$$Z_t^{(\lambda)} = \begin{cases} \dfrac{Z_t^{\lambda} - 1}{\lambda} & , \quad \lambda \neq 0 \\ \ell n\, Z_t & , \quad \lambda \to 0. \end{cases} \qquad (10)$$

La vertu de cette définition est que la limite dans le cas où $\lambda \to 0$ est $\ell n\, Z_t$ (ceci peut être démontré par la règle de l'Hôpital) de sorte que la transformation est continue à $\lambda = 0$ (Box et Cox, 1964, pp.214-215).

4.- La transformation Box-Cox est un cas spécial de la transformation Box-Tukey, laquelle inclut un paramètre de localisation μ. Ainsi, nous avons

$$(Z_t + \mu)^{(\lambda)} = \begin{cases} \dfrac{(Z_t + \mu)^{\lambda} - 1}{\lambda} & , \quad \lambda \neq 0, \quad Z_t + \mu > 0 \\ \ell n\, (Z_t + \mu) & , \quad \lambda \to 0, \quad Z_t + \mu > 0. \end{cases}$$

Lorsque $\lambda=1$, la forme fonctionnelle résultante est linéaire; par ailleurs lorsque $\lambda=-1$, celle-ci est réciproque. Dans le cas où $\lambda=0$, la forme fonctionnelle s'identifie alors comme la forme logarithmique.

L'inverse de la transformation de puissance directe est

$$Z_t = \begin{cases} (\lambda Z_t^{(\lambda)} + 1)^{1/\lambda} & , \quad \lambda \neq 0 \\ \exp(Z_t^{(\lambda)}) & , \quad \lambda \rightarrow 0. \end{cases} \tag{11}$$

L'introduction systématique dans la littérature des transformations de puissance directe sans joindre l'examen de l'impact de ces transformations sur la forme de la partie aléatoire est relativement récente. Parmi les premiers auteurs à considérer des modèles de régression (modèles constitués d'une seule équation) à forme fonctionnelle flexible (et matrice variance-covariance scalaire), nous retrouvons Heckman et Polachek (1974), Kau et Sirmans, Welland (1976, sources citées dans Gaudry et Dagenais, 1978), Spitzer (1976, 1977). Ce dernier a également spécifié un modèle à équations simultanées. Zarembka (1968, 1974) quant à lui, il a posé l'hypothèse voulant que les erreurs suivent un schéma Markovien du premier ordre mais il n'a pas estimé son modèle en considérant simultanément la structure d'autocorrélation et la forme de la partie fixe du modèle.

Jusqu'à présent, notre travail s'est limité à une étude systématique de la partie aléatoire et de la formulation générale de la partie fixe du modèle. Dans les étapes qui suivent, nous allons étendre notre analyse à une spécification conjointe des parties fixe et stochastique lorsque nous augmentons la sophistication de la partie aléatoire du modèle, c'est-à-dire $E(uu') \neq \psi^2 I$.

1.3. Etude systématique de la formulation générale de la partie fixe et de

la partie aléatoire du modèle.

Jusqu'à ce jour, l'analyse systématique Box-Cox de la structure de la partie aléatoire d'un modèle simultanément avec celle de la forme fonctionnelle de sa partie fixe n'a pas fait l'objet de nombreux développements même si la transformation Box-Cox est source de difficultés. Si nous faisons l'esquisse des ouvrages qui mettent en exergue ce problème de la détermination conjointe, nous pouvons noter que Savin et White (1978) furent les premiers auteurs à s'intéresser aux effets de la forme fonctionnelle sur la distribution des erreurs, en considérant l'autocorrélation du premier ordre. Ensuite, nous retrouvons la contribution de Gaudry et Wills (1978), qui ont étudié davantage le problème en incorporant dans le modèle un processus autorégressif d'un ordre supérieur au premier.

Une difficulté inhérente à la transformation Box-Cox est que le terme stochastique est supposé avoir une variance constante à travers la période échantillonnale, alors que cette situation n'est pas toujours irréfutablement omniprésente. La présence de variables transformées dans un modèle non-linéaire peut effectivement engendrer de l'hétéroscédasticité dans le terme d'erreur[5] et ce, en modifiant le rapport des variances. Toutefois, les paramètres de puissance peuvent également compenser pour une mauvaise forme de la partie stochastique du modèle en changeant le caractère hétéroscédastique des erreurs résiduelles. (L'estimation de la forme fonctionnelle par le maximum de vraisemblance peut avoir la propriété de régulariser la distribution du

---

5.- Règle générale, c'est dans les études en coupes instantanées que nous retrouvons le phénomène de l'hétéroscédasticité. Au contraire, dans les études en séries chronologiques, nous postulons si nous nous référons à la théorie classique appliquée au cas linéaire, l'homoscédasticité et l'autocorrélation des résidus. Par ailleurs, lorsque nous utilisons l'approche non-linéaire dans les séries chronologiques, la présence de variables élevées à une certaine puissance λ différente de l'unité peut introduire l'hétéroscédasticité dans les résidus. (Si $\lambda=2$ par exemple, le rapport des variances du grand échantillon sur le petit échantillon va augmenter.)

terme d'erreur (Gaudry et Wills, 1978, p.32)).

Cependant, même si les paramètres de puissance peuvent être plus sensibles à la présence de l'hétéroscédasticité qu'à celle de l'autocorrélation, très peu d'études ont examiné le problème de la détermination conjointe de l'hétéroscédasticité et de la forme fonctionnelle de la partie fixe.

Zarembka (1974) fut le premier auteur à mettre en relief ce problème. Celui-ci a énoncé une mesure[6] approximative de la sensibilité de la forme fonctionnelle à la présence éventuelle d'une forme particulière d'hétéroscédasticité.

Major (1978) a lui aussi apporté une contribution théorique en présentant une nouvelle formulation de l'hétéroscédasticité[7] qu'il a estimé simultanément avec la forme de la partie fixe du modèle.

Une troisième spécification plus générale de l'hétéroscédasticité fut élaborée par Gaudry et Dagenais (1978). Ces derniers ont montré que l'hétéroscédasticité peut être analysée conjointement avec la forme de la partie fixe du modèle. De plus, ils ont introduit une structure d'autocorrélation des erreurs d'ordre multiple.

En plus des difficultés rattachées à la matrice de variance-covariance des erreurs, une autre difficulté à laquelle nous sommes également confrontés suite à la transformation des variables originales par les paramètres de puissance Box-Cox, est celle de la non-normalité des résidus. En effet, si en

---

6.- Cette mesure est développée d'une manière exhaustive dans le texte de Zarembka, p.34.

7.- La formulation de Major s'exprime de la façon suivante:

$$E(u_t^2) = \sigma^2 \exp(\delta_1 X_{1_t} + \delta_2 X_{2_t} + \ldots + \delta_k X_{k_t})\, v$$

où $E(v) = 0$ et $E(v^2) =$ constante.

théorie la procédure Box-Cox est supposée normaliser la distribution des résidus, en pratique cet objectif n'est pas toujours réalisé.

Dans la suite de notre exposé, nous allons utiliser la formulation de l'hétéroscédasticité telle que définie par Gaudry et Dagenais.

1.3.1. Formulation de l'hétéroscédasticité dans un modèle linéaire.

1.3.1.1. Une forme générale pour les erreurs hétéroscédastiques.

Supposons que la matrice variance-covariance des u d'un modèle linéaire s'apparente à la forme suivante:

$$E(uu') = \psi^2 \begin{bmatrix} f(X_1) & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & f(X_t) & 0 \\ 0 & \cdots & 0 & f(X_n) \end{bmatrix} = \Omega \qquad (12)$$

dans laquelle

$$f(x_t) = \{\lambda_u [\delta_o + \sum_m \delta_m X_{m_t}^{(\lambda_{em})}] + 1\}^{1/\lambda_u} > 0 \text{ si } \lambda_u \neq 0. \qquad (13)$$

Les divers éléments de cette expression[8] s'expliquent ainsi:

$X_{m_t}^{(\lambda_{em})}$ désigne la variable X_{m_t} élevée à la puissance λ_{em},

X_{m_t} représente la variable explicative qui est susceptible d'engendrer de l'hétéroscédasticité parmi les k variables indépendantes à l'observation t, pour t=1,....,n,

δ_m est un élément typique du vecteur δ,

λ_{em} est un élément typique du vecteur λ_e,

δ_o et λ_u sont des scalaires.

8.- Nous pouvons obtenir une spécification plus générale de cette équation si, au lieu de choisir un ensemble de variables explicatives ($X_1,\ldots,X_M$) de la partie fixe, nous choisissons un autre ensemble de variables exogènes ($Z_1,\ldots$ Z_M) lesquelles Z_m variables peuvent différer des X_k variables de la partie fixe

La relation (13) laisse voir que les transformations monotoniques Box-Cox de puissance directe ont été appliquées aux X_m variables explicatives de la variance de l'erreur et qu'une transformation de puissance inverse est appliquée aux éléments des parenthèses carrées. La présence des constantes δ_o et ψ est justifiée dans une telle formulation dans le but de préserver l'invariance de la transformation aux unités de mesure des X et de y.

L'équation (13) peut prendre plusieurs formes selon les différentes valeurs que l'on attribue aux paramètres λ_u et λ_{em}.

Si nous posons $\lambda_u = 1$, la relation précitée devient, suite à une multiplication par le terme constant ψ^2, égale à:

$$\tilde{\omega}_{tt} = \tilde{\delta}_o + \sum_m \tilde{\delta}_m X_{m_t}^{(\lambda_{em})} \quad , \quad \text{où } \tilde{\delta}_o = \psi^2(\delta_o + 1) \qquad (14)$$
$$\tilde{\delta}_m = \psi^2 \delta_m .$$

Cette spécification peut inclure plusieurs cas spéciaux: elle inclut la forme linéaire si $\lambda_{em} = 1$ pour tout m et plusieurs formes non-linéaires (racine carrée si $\lambda_{em} = 1/2$ et puissance carrée si $\lambda_{em} = 2$).

Nonobstant, si nous posons $\lambda_u = 0$, l'équation (13) obéit à la condition (11) et elle s'écrit:

$$f(X_t) = \exp\left(\delta_o + \sum_m \delta_m X_{m_t}^{(\lambda_{em})}\right) . \qquad (15)$$

Suite à une multiplication par ψ^2, l'expression (15) devient

$$\tilde{\omega}_{tt} = \psi^2 \exp\left(\delta_o + \sum_m \delta_m X_{m_t}^{(\lambda_{em})}\right) \equiv \sigma^2 \left[\exp\left(\sum_m \delta_m X_{m_t}^{(\lambda_{em})}\right)\right], \qquad (16)$$
$$\text{où } \sigma^2 = \psi^2 e^{\delta_o} .$$

Si $\lambda_{em} = 0$ pour tout m, cette forme se ramène à un modèle multiplicatif multivarié (de type Cobb-Douglas),[9]

9.- A partir de l'équation (16) et de la condition (10), nous obtenons la

$$\hat{\omega}_{tt} = \sigma^2 \prod_m X_{m_t}^{\delta_m} \quad . \tag{17}$$

Cette formulation qui fut développée par Dagum et Dagum (1974) ne répond pas parfaitement au problème d'estimation de l'hétéroscédasticité du terme de l'erreur.

De plus, si nous contraignons δ_m à prendre une valeur égale à 2 pour tout m, la forme (17) devient une forme multiplicative multivariée classique,

$$\hat{\omega}_{tt} = \sigma^2 \prod_m X_{m_t}^{2} \quad . \tag{18}$$

La spécification de Park (1966) est obtenue en posant tous les δ_m de la relation (16) à être égaux à zéro sauf un et en posant le λ_{em} de la variable considérée à être égal à zéro. Ainsi,

$$\hat{\omega}_{tt} = \sigma^2 X_{m_t}^{\delta_m} \quad . \tag{19}$$

Si δ_m est contraint à être égal à deux, nous retrouvons la formulation classique habituelle à savoir

$$\hat{\omega}_{tt} = \sigma^2 X_{m_t}^{2} \quad . \tag{20}$$

Dans la dérivation de la fonction de vraisemblance, nous retiendrons

---

relation (17) de la manière suivante:

$$\hat{\omega}_{tt} = \sigma^2 \left[\exp\left(\sum_m \delta_m \ln X_{m_t}\right)\right]$$

$$= \sigma^2 \left[\exp\left(\delta_1 \ln X_{1_t}\right) \exp\left(\delta_2 \ln X_{2_t}\right) \ldots \exp\left(\delta_m \ln X_{m_t}\right)\right]$$

$$= \sigma^2 \left[\exp\left(\ln X_{1_t}^{\delta_1}\right) \exp\left(\ln X_{2_t}^{\delta_2}\right) \ldots \exp\left(\ln X_{m_t}^{\delta_m}\right)\right]$$

et puisque $\exp(\ln X) = X$, nous avons

$$= \sigma^2 \left[X_{1_t}^{\delta_1}, X_{2_t}^{\delta_2}, \ldots, X_{m_t}^{\delta_m}\right]$$

$$= \sigma^2 \prod_m X_{m_t}^{\delta_m} \quad .$$

la sous-classe (16) au lieu de la sous-classe (14), c'est-à-dire celle où nous contraignons λ_u à égaler 0 au lieu d'égaler 1, car le choix de la formulation (16) permet:

1°) d'éviter des éléments négatifs dans Ω et,

2°) des intéractions entre les variables qui sont les plus susceptibles d'expliquer la variance de l'erreur. (Gaudry et Dagenais, 1978, p.6).

1.3.1.2. Fonction de vraisemblance pour un modèle linéaire.

La fonction de vraisemblance pour le modèle $y = X\beta + u$, lorsque u est distribué normalement avec un vecteur d'espérances mathématiques nulles et une matrice de variance-covariance (12) dans laquelle les termes de la matrice diagonale sont déterminés comme dans (15), peut s'écrire:

$$\mathcal{L}(y;\ \beta,\ \sigma^2,\ \lambda_e,\ \delta) = \left(\frac{1}{\sqrt{2\pi}}\right)^n |\tilde{\Omega}|^{-1/2} \exp\left\{ -\frac{1}{2}(y - X\beta)'\ \tilde{\Omega}^{-1}\ (y - X\beta) \right\} \cdot \quad (21)$$

Si nous substituons la forme structurelle ($y_t = \sum_k \beta_k X_{k_t} + u_t$) à la forme matricielle du modèle et si nous spécifions la valeur de Ω (équation 16) alors, l'équation (21) devient

$$\mathcal{L}(y;\ \beta_k,\ \sigma^2,\ \lambda_{em},\ \delta_m) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi}\ \sigma\ \sqrt{\exp(\sum_m \delta_m X_{m_t})^{(\lambda_{em})}}} \exp\left\{ -\frac{(y_t - \sum_k \beta_k X_{k_t})^2}{2\sigma^2\ \exp(\sum_m \delta_m X_{m_t})^{(\lambda_{em})}} \right\} \cdot \quad (22)$$

Cette fonction de vraisemblance est maximisée par rapport aux β_k de la partie systématique, à σ et aux paramètres inclus dans $f(X_t)$.

La procédure d'estimation utilisée est exposée dans un chapitre ultérieur.

1.3.1.3. Fonction de vraisemblance pour un modèle non-linéaire.

Si la forme générale Box-Cox (9) remplace la forme linéaire spécifiée dans 1.3.1.2. et que la forme des erreurs hétéroscédastiques demeure analogue à la forme (15), alors la fonction de vraisemblance des observations originales y devient

$$\mathcal{L}(y;\ \beta,\ \sigma^2,\ \lambda_y,\ \lambda_x,\ \lambda_e,\ \delta) = \left(\frac{1}{\sqrt{2\pi}}\right)^n |\tilde{\Omega}|^{-1/2} \cdot$$

$$\exp\left\{ -\frac{1}{2} (y^{(\lambda_y)} - X^{(\lambda_x)}\beta)' \tilde{\Omega}^{-1} (y^{(\lambda_y)} - X^{(\lambda_x)}\beta) \right\} |J(\lambda_y;y)| , \qquad (23)$$

où n est le nombre d'observations et,

où $|J(\lambda_y;y)|$ est la valeur absolue du Jacobien de la transformation de $y_t^{(\lambda_y)}$ aux valeurs observées de y_t, ou bien

$$|J(\lambda_y;y)| = \left|\det \frac{\partial y_t^{(\lambda_y)}}{\partial y_t}\right| = \prod_{t=1}^{n} y_t^{\lambda_y - 1} . \qquad (24)$$

Lorsque λ est contraint à la valeur 1 (c'est-à-dire dans un cas de linéarité), le Jacobien est égal à l'unité.

L'équation (23) peut être formulée comme étant:

$$\mathcal{L}(y;\ \beta_k,\ \sigma^2,\ \lambda_y,\ \lambda_{xk},\ \lambda_{em},\ \delta_m) =$$

$$\prod_{t=1}^{n} \frac{y_t^{\lambda_y - 1}}{\sqrt{2\pi}\ \sigma \sqrt{\exp(\Sigma_m \delta_m X_{m_t}^{(\lambda_{em})})}} \exp\left\{ -\frac{(y_t^{(\lambda_y)} - \Sigma_k \beta_k X_{k_t}^{(\lambda_{xk})})^2}{2\sigma^2 \exp(\Sigma_m \delta_m X_{m_t}^{(\lambda_{em})})} \right\} . \qquad (25)$$

Comme nous pouvons le voir, cette équation incorpore simultanément la forme de l'hétéroscédasticité et la forme de la partie fixe du modèle.

Dans la prochaine section, nous allons énoncer une formulation complète du modèle non-linéaire. Pour arriver à une telle spécification, nous allons modifier la nature de la matrice Ω en incorporant une structure d'autocorrélation multiple aux erreurs hétéroscédastiques. Nonobstant, avant d'étendre notre analyse à un schéma autorégressif multiple, nous allons pour simplifier le problème, considérer le processus autorégressif du premier ordre.

1.3.2. Formulation de l'hétéroscédasticité et de l'autocorrélation du premier ordre dans un modèle non-linéaire.

La relaxation de l'hypothèse d'indépendance temporelle entre les erreurs résiduelles nous incite à reformuler le problème de la manière suivante:

$$y_t^{(\lambda_y)} = \sum_k \beta_k X_{k_t}^{(\lambda_{xk})} + u_t \quad , \tag{26}$$

avec

$$u_t = (f(X_t))^{1/2} v_t \quad , \tag{27}$$

où $E(v_t v_{t'}) = \psi^2$, pour $t=t'$

$\neq 0$, pour $t \neq t'$.

(Si $t'=t-1$, alors $E(v_t v_{t-1}) = \rho_1 \psi^2$).

Puisque nous supposons que la variable aléatoire v_t suit un processus stationnaire du premier ordre, nous pouvons poser

$$v_t = \rho_1 v_{t-1} + s_t \quad , \tag{28}$$

où $|\rho_1| < 1$ et,

où s_t est normalement et indépendamment distribué avec moyenne zéro,

$E(s_t) = 0$ et matrice variance-covariance,

$E(s_t s_{t'}) = E(s_t^2) = \psi^{*2} = \psi^2 (1 - \rho_1^2)$, pour $t=t'$

$= 0$, pour $t \neq t'$.

D'après Prais et Winsten, Kadiyala (1968, p.94), Goldfeld et Quandt (1972, p.184) et certains autres, si les séries sont stationnaires, le vecteur d'erreurs s peut être défini comme

$$s = Rv \tag{29}$$

où la matrice (n × n) de transformation R s'exprime comme

$$R = \begin{bmatrix} \sqrt{1-\rho_1^2} & 0 & \dots & 0 & 0 \\ -\rho_1 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 1 & 0 \\ 0 & 0 & \dots & -\rho_1 & 1 \end{bmatrix} . \tag{30}$$

A partir de (29) et (30), nous pouvons écrire que $s_1 = \sqrt{1 - \rho_1^2}\, v_1$.

Considérant la transformation précitée qui nous permet de transformer l'équation (26) en un modèle de régression standard, nous pouvons postuler que:

$$u = \text{diag}\,[(f(X_t))^{1/2}]R^{-1}s,$$

ou encore

$$u = HR^{-1}s,$$

avec $E(uu') = E(HR^{-1}ss'(R^{-1})'H')$

$$= HR^{-1}(R^{-1})'H'E(ss')$$

et puisque $E(ss') = \psi^{*2}$ par hypothèse, nous obtenons

$$E(uu') = HR^{-1}(R^{-1})'H'\psi^{*2}$$

$$= \Omega^* .$$

. <u>Fonction de vraisemblance en présence de l'hétéroscédasticité et de l'autocorrélation du premier ordre</u>.

La prise en considération simultanée de l'hétéroscédasticité et de l'autocorrélation du premier ordre dans le cas particulier où une valeur de

zéro est attribuée au paramètre de puissance λ_u, implique une transformation de la fonction de vraisemblance telle que définie dans (23). Ainsi,

$$\mathcal{L}(y;\ \beta,\ \sigma^{*2},\ \rho_1,\ \lambda_y,\ \lambda_x,\ \lambda_e,\ \delta) = \left(\frac{1}{\sqrt{2\pi}\ \sigma^*}\right)^n |\tilde{H}|^{-1/2}\ |R|\ .$$

$$\exp\left\{-\frac{1}{2}\left(y^{(\lambda_y)} - X^{(\lambda_x)}\beta\right)'\ \tilde{\Omega}^{*-1}\ \left(y^{(\lambda_y)} - X^{(\lambda_x)}\beta\right)\right\}\ |J(\lambda_y;y)|\ , \qquad (31)$$

où $\tilde{\Omega}^* = \tilde{H}R^{-1}(R^{-1})'\tilde{H}'\sigma^{*2}$.

Ceci peut être démontré de la façon suivante:

Puisque H = diagonale $[(f(X_t))^{1/2}]$, si nous posons $\lambda_u = 0$ alors,

$$H = \text{diagonale}\ [\exp(\delta_o + \Sigma_m \delta_m X_{m_t}^{(\lambda_{em})}/2)]\ .$$

En multipliant cette expression par ψ^{*2} nous obtenons,

$$\psi^{*2} H = \psi^{*2}\ \text{diagonale}\ [\exp(\delta_o + \Sigma_m \delta_m X_{m_t}^{(\lambda_{em})}/2]$$

$$= \psi^2\ (1 - \rho_1^2)\ \exp \delta_o\ \text{diagonale}\ [\exp(\Sigma_m \delta_m X_{m_t}^{(\lambda_{em})}/2)]\ .$$

Comme $\psi^2 \exp \delta_o = \sigma^2$ (tel que noté précédemment), nous pouvons écrire que

$$\psi^{*2} H = \sigma^2\ (1 - \rho_1^2)\ \text{diagonale}\ [\exp(\Sigma_m \delta_m X_{m_t}^{(\lambda_{em})}/2)]$$

$$= \sigma^{*2}\ \tilde{H}\ .$$

Puisque $|R| = (1 - \rho_1^2)^{1/2}$, l'expression (31) peut se ramener par homologie à

$$\mathcal{L}(y;\ \beta_k,\ \sigma^{*2},\ \rho_1,\ \lambda_y,\ \lambda_{xk},\ \lambda_{em},\ \delta_m) = (1 - \rho_1^2)^{1/2}\ \frac{y_1^{\lambda_y - 1}}{\sqrt{2\pi}\ \sigma^*\ \sqrt{\exp(\Sigma_m \delta_m X_{m_1}^{(\lambda_{em})})}}\ .$$

$$\exp\left\{ - \frac{(1 - \rho_1^2)\,(y_1^{(\lambda_y)} - \sum_k \beta_k X_{k_1}^{(\lambda_{xk})})^2}{2\sigma^{*2}\,\exp(\sum_m \delta_m X_{m_1}^{(\lambda_{em})})} \right\} \cdot \prod_{t=2}^{n} \frac{y_t^{\lambda_y - 1}}{\sqrt{2\pi}\,\sigma^* \sqrt{\exp(\sum_m \delta_m X_{m_t}^{(\lambda_{em})})}} \cdot$$

$$\exp\left\{ - \frac{1}{2\sigma^{*2}} \left[\frac{(y_t^{(\lambda_y)} - \sum_k \beta_k X_{k_t}^{(\lambda_{xk})})}{\sqrt{\exp(\sum_m \delta_m X_{m_t}^{(\lambda_{em})})}} - \rho_1 \frac{(y_{t-1}^{(\lambda_y)} - \sum_k \beta_k X_{k_{t-1}}^{(\lambda_{xk})})}{\sqrt{\exp(\sum_m \delta_m X_{m_{t-1}}^{(\lambda_{em})})}} \right]^2 \right\} . \quad (32)$$

La différence entre les relations (25) et (32) réside en ce que les termes qui se réfèrent à la première observation se distinguent des autres.

1.3.2.1. Formulation simplifiée en présence de l'autocorrélation du premier ordre.

La présence de la première observation dans un schéma autorégressif du premier ordre peut engendrer diverses complications. Ainsi, si $|\rho_1|$ tend vers l'unité, alors $(1 - \rho_1^2)^{1/2}$ tendra vers zéro, et la fonction de vraisemblance convergera vers zéro.

Nonobstant, la complexité de la spécification (32) peut être simplifiée par l'omission de l'observation initiale. D'après Beach et MacKinnon (1978, p.52), le fait de ne pas introduire la première observation n'implique aucune différence asymptotiquement. Ces auteurs ont cependant démontré à l'aide d'un modèle linéaire homoscédastique constitué d'une constante et d'une seule variable explicative que l'exclusion de cette observation dans le cas d'un faible échantillonnage peut occasionner une perte d'efficacité[10] (v.g.

10.- Utilisant une analyse de Monte-Carlo (pp.54-56), ils ont réussi à démontrer pour des échantillons constitués de 20 et de 50 observations et pour différentes valeurs de ρ, que des différences substantielles peuvent être notées selon qu'ils utilisent la procédure conventionnelle de Cochrane-Orcutt qui maximise les seules n-1 observations ou selon qu'ils utilisent une procédure plus sophistiquée (Full Maximum Likelihood) qui maximise les n observa-

dans le biais et dans la racine de la moyenne des erreurs au carré).

Sur des bases purement théoriques, Beach et MacKinnon soulignent qu'il serait préférable d'utiliser une procédure du maximum de vraisemblance qui ne tienne pas compte de la première observation mais qui incorpore en même temps à priori la restriction de stationnarité $|\rho| < 1$.

En conséquence, si nous respectons les exigences d'une approche "de préférence" (telle que définie par les deux auteurs précités), la première ligne de la matrice R (30) disparaît, tout comme le premier terme du Jacobien et le terme pour u_1^2 dans (32). La fonction de vraisemblance (32) devient alors

$$\mathcal{L}(y;\ \beta_k,\ \sigma^{*2},\ \rho_1,\ \lambda_y,\ \lambda_{xk},\ \lambda_{em},\ \delta_m) =$$

$$\prod_{t=2}^{n} \frac{[y_t^{\lambda_y - 1}]^{\sim}}{\sqrt{2\pi}\ \sigma^*} \exp\left\{ - \frac{1}{2\sigma^{*2}} \left([y_t^{(\lambda_y)}]^{\tilde{}\star} - \sum_k \beta_k [X_{k_t}^{(\lambda_{xk})}]^{\tilde{}\star}\right)^2 \right\} \qquad (33)$$

avec

$$[\phi_t]^{\sim} = \frac{\phi_t}{\sqrt{\exp\left(\sum_m \delta_m X_{m_t}^{(\lambda_{em})}\right)}} \quad \text{et} \quad [\phi_t]^{\tilde{}\star} = \frac{\phi_t}{\sqrt{\exp\left(\sum_m \delta_m X_{m_t}^{(\lambda_{em})}\right)}} - \rho_1 \frac{\phi_{t-1}}{\sqrt{\exp\left(\sum_m \delta_m X_{m_{t-1}}^{(\lambda_{em})}\right)}}. \qquad (34)$$

1.3.2.2. Formulation simplifiée en présence de l'autocorrélation multiple.

Cette simplification de la fonction de vraisemblance peut être étendue

---

tions. Les résultats obtenus ne vérifient pas leurs anticipations car il y a virtuellement une plus grande différence dans le gain d'efficacité d'inclure la première observation dans l'échantillon de 50 observations que dans l'échantillon de 20 observations. D'après eux, l'équivalence asymptotique des procédures conventionnelles et complètes pour atteindre une efficacité égale ne peut être assurée que par l'utilisation d'un échantillon de très grande importance.

à une structure autorégressive stationnaire multiple ($\rho = \rho_1, \ldots, \rho_\ell, \ldots, \rho_r$) en éliminant autant d'observations que l'ordre d'autocorrélation le plus élevé le requiert. L'emploi de la formulation simplifiée de la fonction de vraisemblance permet d'éluder les problèmes de calcul considérables qu'impliquerait la formulation complète de la fonction de vraisemblance.

Dans le cas où le nombre d'observations ignorées est égal à r, les expressions (33) et (34) deviennent respectivement

$$\mathcal{L}(y;\ \beta_k,\ \sigma^{**2},\ \rho_\ell,\ \lambda_y,\ \lambda_{xk},\ \lambda_{em},\ \delta_m) =$$

$$\prod_{t=1+r}^{n} \frac{[y_t^{\lambda_y - 1}]\sim}{\sqrt{2\pi}\ \sigma^{**}} \exp\left\{ -\frac{1}{2\sigma^{**2}} \left([y_t^{(\lambda_y)}]\tilde{}^{**} - \Sigma_k \beta_k [X_{k_t}^{(\lambda_{xk})}]\tilde{}^{**}\right)^2 \right\} \qquad (35)$$

avec $[\phi_t]\sim$ caractérisé comme dans (34), et $\sigma^{**} > 0$

$$\text{et} \quad [\phi_t]\tilde{}^{**} = \frac{\phi_t}{\sqrt{\exp(\Sigma_m \delta_m X_{m_t}^{(\lambda_{em})})}} - \sum_{\ell=1}^{r} \rho_\ell \frac{\phi_{t-\ell}}{\sqrt{\exp(\Sigma_m \delta_m X_{m_{t-\ell}}^{(\lambda_{em})})}} . \qquad (36)$$

Le logarithme de la fonction de vraisemblance peut être formulé comme étant:

$$L(y;\ \beta_k,\ \sigma^{**2},\ \rho_\ell,\ \lambda_y,\ \lambda_{xk},\ \lambda_{em},\ \delta_m) =$$

$$L^* = (\lambda_y - 1) \sum_{t=1+r}^{n} \ell n\ y_t - \left(\frac{n-r}{2}\right) \ell n\ \sigma^{**2} - \frac{1}{2} \sum_{t=1+r}^{n} \ell n\ f(X_t)$$

$$- \frac{1}{2\sigma^{**2}} \sum_{t=1+r}^{n} [y_t^{**} - \Sigma_k \beta_k X_{k_t}^{**}]^2 \qquad (37)$$

où

$$1)\ y_t^{**} = y_t^{*} - \sum_{\ell} \rho_\ell\, y_{t-\ell}^{*} \tag{38}$$

$$\cdot\ y_t^{*} = \frac{y_t^{(\lambda_y)}}{\sqrt{f(X_t)}} \tag{39}$$

$$2)\ X_{k_t}^{**} = X_{k_t}^{*} - \sum_{\ell} \rho_\ell\, X_{k_{t-\ell}}^{*} \tag{40}$$

$$\cdot\ X_{k_t}^{*} = \frac{X_{k_t}^{(\lambda_{xk})}}{\sqrt{f(X_t)}}\ . \tag{41}$$

CHAPITRE II

PROCEDURE D'ESTIMATION.

L'optimisation d'une fonction de vraisemblance d'un modèle non-linéaire nécessite l'utilisation d'un algorithme qui maximise une fonction appropriée ou qui minimise une somme d'erreurs au carré.

La technique de maximisation suggère ici de maximiser le logarithme de la fonction de vraisemblance concentrée afin de minimiser le nombre de paramètres impliqués dans la maximisation de (37). Ainsi, si nous concentrons le logarithme de la fonction de vraisemblance sur σ^{**2} et alors sur les β_k, les seuls paramètres de la fonction qui seront intégrés dans le processus de maximisation seront[1] ρ, λ_y, λ_x, λ_e et δ.

Le logarithme de la fonction de vraisemblance concentrée peut donc être écrit comme[2]

$$L(y;\ \rho,\ \lambda_y,\ \lambda_x,\ \lambda_e,\ \delta) =$$

$$L^{***} = (\lambda_y - 1) \sum_{t=1+r}^{n} \ell n\ y_t - \left(\frac{n-r}{2}\right) \ell n\ \hat{\sigma}^{**^2}(\rho, \lambda_y, \lambda_x, \lambda_e, \delta)$$

$$- \frac{1}{2} \sum_{t=1+r}^{n} \ell n\ f(X_t) \qquad (42)$$

où $\hat{\sigma}^{**^2}(\rho,\ \lambda_y,\ \lambda_x,\ \lambda_e,\ \delta) =$

1.- Les indices ℓ, k, m sont supprimés respectivement pour ρ, λ_x, λ_e et δ afin d'exprimer ces paramètres sous une forme vectorielle.

2.- La démonstration des principales étapes pour obtenir L^{***} est présentée dans l'appendice A.

$$\frac{1}{(n-r)} [(y^{**} - X^{**}\hat{\beta}(\rho,\lambda_y,\lambda_x,\lambda_e,\delta))' (y^{**} - X^{**}\hat{\beta}(\rho,\lambda_y,\lambda_x,\lambda_e,\delta))] , \quad (43)$$

$$\hat{\beta}(\rho, \lambda_y, \lambda_x, \lambda_e, \delta) = (X^{**'}X^{**})^{-1} X^{**'} y^{**} . \quad (44)$$

Puisque la fonction est concentrée sur σ^{**2}, les β qui maximisent L^{***} pour des valeurs données de ρ, λ_y, λ_x, λ_e et δ sont les estimés des moindres carrés ordinaires.

Nonobstant, si les estimés de β sont calculés de la même façon que ceux des moindres carrés ordinaires, ils possèdent toutefois les propriétés asymptotiques des estimateurs du maximum de vraisemblance à savoir la convergence et l'efficacité asymptotique.

La maximisation [3] du logarithme de la fonction de vraisemblance (42) requiert l'utilisation d'un algorithme qui maximise une fonction non-linéaire par rapport à ρ, λ_y, λ_x, λ_e, δ. Dans le cas précis qui nous préoccupe, l'algorithme de résolution du maximum de vraisemblance est inféodé à l'algorithme de Fletcher et Powell (1963).

La méthode décrite par ces deux auteurs est applicable au cas d'un gradient défini.

Cependant, puisque l'algorithme considéré minimise la fonction de vraisemblance, nous devons renverser les signes de la fonction et de ses dérivées: min - (L^{***}) = max (L^{***}).

Si la fonction est continue, la condition suffisante pour atteindre

---

3.- La solution du maximum de vraisemblance de l'équation (42) est obtenue à partir du programme DSDM (Demande Spatiale Directe par Mode). Ce programme qui fut conçu par Tran Cong Liem (1979) permet de résoudre des problèmes ayant un grand nombre de paramètres. De nature flexible, ce programme sophistiqué nous permet de fixer ou encore de laisser estimer les différents paramètres du modèle.

un maximum (au moins local) est que les dérivées premières partielles définies analytiquement soient égales à zéro. Cependant, étant donné que le gradient n'est pas invariant aux unités de mesure, il est possible que les dérivées premières ne s'annulent pas numériquement; la solution finale correspondante est également celle d'un maximum local.

Si nous possédons l'information requise pour nous assurer d'un maximum local, il n'existe par ailleurs aucune approche généralisée pour nous certifier que le maximum local obtenu est équivalent à un maximum global.

La résolution d'un maximum global exige qu'à partir de toutes les solutions de départ d'un certain espace[4] défini raisonnablement, la fonction converge vers un point unique. Si chaque solution initiale (de départ) restreinte au même espace n'assure pas l'unicité de la solution finale, alors le maximum global sera celui qui parmi tous les maxima locaux garantit la valeur la plus élevée de la fonction de vraisemblance.

Quand la convergence de la fonction est assurée, c'est-à-dire lorsqu'un maximum (au moins local) de L^{***} est indubitablement atteint, il est alors possible de calculer la matrice de covariance asymptotique pour les différents paramètres[5] de la partie fixe et de la partie stochastique du modèle. Le calcul d'une telle matrice suggère l'utilisation de la relation statistique

---

4.- Dans notre cas, pour maximiser sur l'espace entier des paramètres, il faudrait choisir des valeurs de λ comprises dans la région suivante: $-20 < \lambda < 20$. Ces valeurs sont également choisies dans le but d'éviter des problèmes de calcul numérique.
Considérant les coûts (temps et monétaire) qu'impliquerait la localisation d'un maximum global, nous n'allons pas opérer dans ce sens.

5.- Il est à noter que la procédure utilisée pour le calcul des écarts-types des divers paramètres est approximative et inexacte. En effet, lors du calcul de la matrice de covariance asymptotique des paramètres, la covariance entre σ^{**2} et les autres paramètres (β, ρ, λ_y, λ_x, λ_e, δ) est considérée comme étant nulle de sorte que l'expression (46) définie dans la page 28, ne s'avère pas tout à fait exacte.

fondamentale que "la matrice de covariance asymptotique de l'estimateur du maximum de vraisemblance est égale à l'inverse de la matrice de covariance du gradient de la fonction de vraisemblance", (Kendall et Stuart, 1967, p.55).

Le gradient de la fonction de vraisemblance logarithmique concentrée sur σ^{**2} (le calcul de cette dernière est présenté dans l'annexe A v.g. équation (A-6)) peut s'écrire comme suit:

$$\frac{\partial L^{**}}{\partial S} = \sum_{t=1+r}^{n} D_t \tag{45}$$

où S est un vecteur colonne de M éléments contenant les paramètres β, ρ, λ_y, λ_x, λ_e et δ,

et D_t est un vecteur colonne de M éléments.

Ce vecteur est défini ainsi:

$$D_t = \frac{\partial (\lambda_y - 1) \ln y_t}{\partial S} - \frac{(y_t^{**} - \sum_k \beta_k X_{k_t}^{**})}{\hat{\sigma}^{**2}} \frac{\partial (y_t^{**} - \sum_k \beta_k X_{k_t}^{**})}{\partial S}$$

$$- \frac{1}{2} \frac{\partial \ln f(X_t)}{\partial S} . \tag{46}$$

Au maximum de la fonction de vraisemblance logarithmique, la matrice de covariance asymptotique M × M pour $S = \hat{S}$ est égale à la somme du produit croisé des dérivées premières,

$$COV(\hat{S}) = \left[\sum_{t=1+r}^{n} D_t D'_t \right]^{-1}_{S=\hat{S}} . \tag{47}$$

Les différentes dérivées impliquées dans l'expression D_t sont présentées dans l'article de Tran Cong Liem (1979, pp.19-22).

* * *

Dans le présent rapport de recherche, la fonction de vraisemblance (42) sera maximisée selon une série d'hypothèses[6] (contraignantes et non contraignantes) émises à priori sur les différents paramètres du modèle.

La détermination du modèle exprimé dans une forme générale s'obtient par la combinaison des équations (22), (23) et (24)[7]. Ainsi suite à cette juxtaposition, nous obtenons le modèle général suivant:

$$\frac{y_t^{(\lambda_y)}}{\sqrt{\exp(\Sigma_m \delta_m X_{m_t}^{(\lambda_{em})})}} - \sum_{\ell=1}^{r} \rho_\ell \frac{y_{t-\ell}^{(\lambda_y)}}{\sqrt{\exp(\Sigma_m \delta_m X_{m_{t-\ell}}^{(\lambda_{em})})}} =$$

$$\Sigma_k \beta_k \frac{X_{k_t}^{(\lambda_{xk})}}{\sqrt{\exp(\Sigma_m \delta_m X_{m_t}^{(\lambda_{em})})}} - \Sigma_k \sum_{\ell=1}^{r} \beta_k \rho_\ell \frac{X_{k_{t-\ell}}^{(\lambda_{xk})}}{\sqrt{\exp(\Sigma_m \delta_m X_{m_{t-\ell}}^{(\lambda_{em})})}} + w_t \qquad (48)$$

dans lequel,

$$w_t = v_t - \sum_{\ell=1}^{r} \rho_\ell v_{t-\ell} \quad , \quad \text{où } t=1+r, \ldots, n \ , \quad \ell = 1, \ldots, r \ . \qquad (49)$$

Nous pouvons noter que le terme d'erreur w_t respecte l'hypothèse de normalité à une nuance près. Comme le souligne Ramsey (1971), la distribution de l'erreur ne peut être strictement normale puisqu'une transformation de puissance ne permet pas en général les grandes valeurs négatives pour le terme

---

6.- Les différentes hypothèses exprimées seront énumérées ultérieurement (voir chapitre 3.2).

7.- Etant donné que nous sommes en présence d'une structure d'autocorrélation multiple, nous identifierons le terme résiduel non plus comme étant s_t mais plutôt comme étant w_t. Ce terme d'erreur obéit aux mêmes hypothèses que s_t.

résiduel. (Si w_t était fortement négatif lorsque $\lambda_y \lessgtr 0$, alors y_t pourrait devenir négatif; or, les valeurs négatives sont inadmissibles pour la variable dépendante y_t lorsque celle-ci est sujette à des transformations Box-Cox). Cependant, il ajoute que si la probabilité d'obtenir de telles valeurs est assez faible, la distribution de l'erreur peut être considérée comme approximativement normale.

Pour une discussion plus exhaustive sur l'hypothèse de normalité, nous pouvons consulter l'appendice B.

Pour une valeur donnée du paramètre de puissance λ_u (nous posons $\lambda_u = 0$), l'équation (48) est non-linéaire dans les coefficients ρ_ℓ, λ_y, λ_{xk}, λ_{em} et δ_m. Pour obtenir les estimés de ces paramètres, nous identifions le schéma autorégressif en examinant, dans l'esprit de Box et Jenkins (1970, pp.496-497), les fonctions d'autocorrélation et d'autocorrélation partielle des estimateurs non-biaisés des résidus[8] a_t obtenus en appliquant les moindres carrés ordinaires[9] à l'équation (26) pour $\lambda_y = \lambda_{xk} = 1$, (Gaudry et Wills, 1978, p.53).

La spécification initiale du processus autorégressif étant bien définie, il sera possible par la suite de tester cette spécification simultanément avec la forme fonctionnelle de la partie systématique et de la partie aléatoire du modèle. La détermination de la structure autorégressive demeurera inchangée même si nous émettons à priori différentes hypothèses explicites sur

---

8.- Ici, le terme résiduel u_t est égal à v_t et non pas à $(f(X_t))^{1/2} v_t$ puisque si $\lambda_{em} = \delta_m = 0$ (c'est-à-dire postulat d'homoscédasticité), la composante $(f(X_t))^{1/2} = 1$.

9.- Les moindres carrés ordinaires que nous considérons ici, sont en fait un pseudo-moindres carrés ordinaires puisqu'une constante équivalente à l'unité est soustraite de la variable Z. En effet, $Z^{(\lambda)} = \frac{Z^{(\lambda)} - 1}{Z}$ pour $\lambda = 1$.

les paramètres[10]. Cette situation existe non pas parce que la structure d'autocorrélation est indépendante de la forme fonctionnelle mais plutôt parce qu'il n'y aurait aucun sens de modifier constamment le schéma d'autocorrélation des erreurs résiduelles. En effet, un des objectifs de notre étude n'est pas de spécifier une structure d'autocorrélation différente pour les diverses valeurs des paramètres inclus dans les vecteurs de transformation λ_y, λ_x, λ_e et δ mais de voir l'impact de ces paramètres sur les estimateurs d'une structure de paramètres β_k donnée.

Les estimés de β (ainsi que de σ^{**2}) sont calculés à partir des valeurs finales de $\hat{\rho}$, $\hat{\lambda}$ et $\hat{\delta}$.

10.- Néanmoins, pour vérifier si le schéma autorégressif des différentes fonctions à maximiser est compatible avec la structure d'autocorrélation des moindres carrés ordinaires, nous allons tout de même considérer les fonctions d'autocorrélation et d'autocorrélation partielle pour le terme résiduel w_t (défini dans l'équation 48) lesquelles sont calculées à l'étape finale de la maximisation de L^{***}.

CHAPITRE III

MODELE A ESTIMER.

3.1. Application à un modèle de demande de transport en commun.

La charpente théorique de la détermination conjointe de la forme fonctionnelle et de la spécification aléatoire du modèle étant clairement conceptualisée et la procédure d'estimation étant bien définie, nous allons maintenant présenter l'esquisse schématique du modèle à estimer.

Le cadre d'analyse adopté pour tenter de résoudre les problèmes qui gouvernent la forme fonctionnelle est assujetti à une fonction de demande de transport en commun en contexte urbain.

La fonction de demande qui est analysée ici cherche à expliquer les flux de voyages des individus (plus particulièrement des écoliers) sur le territoire de la Communauté urbaine de Montréal (pour toutes les paires origines-destinations agrégées) à partir d'un mode de transport spécifique. Les déplacements sont fonctions de variables socio-économiques (variables micro-économiques et variables d'activité), de caractéristiques particulières associées au transport et d'attributs divers.

Les données de base pour les déplacements des écoliers dans le réseau et pour les 26 variables explicatives du modèle considéré (spécification de M. Gaudry, 1973) sont des données mensuelles en séries chronologiques qui correspondent à 181 observations (nulles et non-nulles) consécutives pour la période s'établissant entre décembre 1956 et décembre 1971.

L'équation de demande[1] non séquentielle (avec ajustement instantané) retenue s'exprime de la façon suivante:

$$y_t^d = h^d (P_t, I_t, T_t, C_t, A_t, O_t) \tag{50}$$

dans laquelle

y_t^d : représente le nombre de passagers du transport en commun,

P_t : est un vecteur contenant les prix d'un déplacement effectué soit par le transport public, soit par un mode de transport concurrent (en l'occurrence l'automobile) et le prix des biens de consommation autres que le transport,

I_t : est un terme se référant au revenu,

T_t : est un vecteur renfermant les variables de niveau de service ou encore de temps,

C_t : est un vecteur contenant les variables de confort,

A_t : dénote un vecteur de variables d'activité (école, magasinage, événements spéciaux) qui génèrent les voyages,

O_t : représente un vecteur de variables comptables et d'agrégation.

Pour effectuer l'estimation de l'équation (50), nous devons l'intégrer au modèle général non-linéaire. Ainsi, si nous colligeons cette spécification et les équations (26), (27) et (49), nous obtenons:

$$\begin{aligned} y_t^{(\lambda_y)} = {} & \sum_q \sum_{k=1}^{3} \beta_{qk} P_{qk_t}^{\lambda_{x(1)}} + \beta_4 I_t^{\lambda_{x(2)}} + \sum_q \sum_{k=5}^{7} \beta_{qk} T_{qk_t}^{\lambda_{x(3)}} \\ & + \sum_{k=8}^{11} \beta_k C_{k_t} + \left(\sum_{k=13,17} \beta_k A_{k_t}^{\lambda_{x(4)}} + \sum_{k=12,14} \beta_k A_{k_t} \right) \\ & + \left(\sum_{k=21}^{23} \beta_k O_{k_t}^{\lambda_{x(5)}} + \sum_{k=18,19,20,24}^{26} \beta_k O_{k_t} \right) + u_t \end{aligned} \tag{51}$$

1.- La discussion de l'équilibre du marché et de l'identification de la fonction de demande est exposée dans une des études de Gaudry (1975, pp.6-9).

$$\text{avec } u_t = \sqrt{\exp(\Sigma_m \delta_m X_{m_t}^{(\lambda_{em})})}\ v_t$$

$$v_t = \sum_{\ell=1}^{r} \rho_\ell\ v_{t-\ell} + w_t\ .$$

La structure d'autocorrélation des 1°, 3°, 4° et 12° ordres est considérée. Ainsi t=13 ,...., 181.

Dans le modèle, q représente un indice relié au mode et k est un indice indiquant le rang des variables explicatives impliquées dans la régression. De plus, (1) ,...., (5)=h et où h représente le h ième groupe de X_k.

3.1.1. Choix de la paramétrisation de la partie systématique du modèle.

Dans ce modèle, nous n'utilisons pas la transformation Box-Cox "pure" car certaines des variables ne peuvent être soumises à une transformation.

En effet, l'utilisation de la procédure Box-Cox requiert des observations positives pour chacune des variables (dépendante et indépendante); or, dans la spécification de notre modèle, nous retraçons trois groupes de variables explicatives parmi lesquelles seules les variables appartenant au premier groupe peuvent subir des transformations paramétriques.

Ce premier groupe est composé par des variables uniquement positives, c'est-à-dire 1°) $X > 0$.

Les deux autres groupes qui ne peuvent être paramétrisés sont constitués respectivement par des variables "semi-dichotomiques" et par des variables "dichotomiques",

c'est-à-dire 2°) $\quad S\text{-}D \begin{cases} S\text{-}D_i \Rightarrow X > 0 \\ \qquad\qquad \text{où } i + j = t \\ S\text{-}D_j \Rightarrow X = 0 \end{cases}$

$$3^{\circ}) \quad D \begin{cases} D_i \Rightarrow X = 1 \\ \qquad\qquad \text{où } i + j = t \\ D_j \Rightarrow X = 0 \, . \end{cases}$$

Afin de discerner quelles sont les variables qui sont sujettes à des transformations Box-Cox, nous allons dresser la liste des variables[2] (voir tableau 1) en indiquant le nombre d'observations positives.

Le tableau 1 montre qu'il y a 12 variables indépendantes qui peuvent être transformées. Les éléments des vecteurs P_t, T_t et du scalaire I_t peuvent être tous paramétrisés. Par ailleurs, seulement 2 des 6 variables d'activité et 3 des 9 caractéristiques diverses peuvent être transformées par la procédure Box-Cox. Des 14 composantes qui ne peuvent subir aucune transformation, il y a les 10 variables restantes pour les vecteurs A_t et O_t et les 4 variables de confort.

Dans notre modèle, nous appliquons un λ_x commun[3] pour chacune des variables (positives) qui composent une catégorie distincte de variables explicatives.

Etant donné qu'il y a seulement 5 des 6 groupes qui possèdent au moins une variable pouvant être transformée, la paramétrisation Box-Cox des variables explicatives de notre modèle nécessitera essentiellement l'utilisation de 5 lambdas différents.

2.- Les définitions des variables explicatives sont données dans l'appendice D.

3.- L'utilisation d'un lambda x commun par groupe de variables (le groupement se fait soit par norme, soit par catégorie) est rendue nécessaire par le fait que le programme DSDM permet de traiter simultanément jusqu'à 6 transformations distinctes de la variable dépendante et de groupes de variables indépendantes.

De plus, il permet l'inclusion de douze paramètres associés à la forme de l'hétéroscédasticité (λ_u, λ_{em}, δ_m) et de cinq paramètres d'autocorrélation (ρ_ℓ).

TABLEAU 1

Nombre d'observations positives pour les différentes variables.

Groupe	Code	Variable	Nbr. obs. positives
Prix (P)	FSCD	1. Tarif réel étudiant	181
	NC	2. Nombre d'automobiles	181
	P	3. Indice des prix à la consommation	181
Revenu (I)	RAWEM	4. Salaire réel moyen	181
Temps (T)	TW	5. Temps d'attente du t.e.c.	181
	TT	6. Temps de transport en t.e.c.	181
	T2T	7. Temps de transport en automobile	181
Confort (C)	TLK	8. Inconfort dû à la température	180
	RFK	9. Inconfort dû à la pluie	179
	SFK	10. Inconfort dû à la neige	96
	CSFMR	11. Inconfort dû à la neige accumulée	95
Activités (A)	RSCMP	12. Indice de présences scolaires	152
	SHONF	13. Indice de déplacements-magasinage	181
	EX7	14. Expo'67	7
	EX8	15. Expo'68	6
	STR	16. Grève à la C.T.C.U.M.	3
	TEND	17. Trend	181
Autres (O)	ANT	18. Ventes anticipées	1
	ANTL1	19. Ventes anticipées retardées	1
	ZSC	20. Double comptage zonal	108
	WD	21. Jours ouvrables/mois	181
	SAT	22. Samedis/mois	181
	SHD	23. Dimanches & jours fériés/mois	181
	CHEAT2	24. Contrôle de la carte d'écolier	15
	STRUCT1	25. Restructuration des lignes	63
	CARDEX	26. Expiration de la carte étudiant	18

3.1.2. Choix de la paramétrisation de la partie stochastique du modèle.

La spécification de la forme fonctionnelle de la partie aléatoire nécessite de distinguer quelle(s) variable(s) sera (seront) impliquée(s) dans la forme de l'hétéroscédasticité.

Nous nous proposons d'estimer la forme de l'hétéroscédasticité en supposant que les variables d'activité ($X_{13} \equiv$ *SHONF* et $X_{17} \equiv$ *TEND*) sont la source de l'hétéroscédasticité. Le choix de ces variables n'est pas établi sur une base arbitraire. En effet, dans les séries chronologiques, la théorie classique suggère pour un modèle linéaire que les variables qui subissent des fluctuations saisonnières sont les plus susceptibles d'engendrer l'hétéroscédasticité des résidus.

Dans un premier temps, la transformation Box-Cox est appliquée à chacune des deux variables prises isolément, puis dans un second temps aux deux variables prises conjointement. Dans le dernier cas, nous posons un λ_e différent pour chacun des facteurs considérés comme étant la source de l'hétéroscédasticité.

Dans cette étude, deux formes d'hétéroscédasticité sont estimées. En plus de la forme générale de l'hétéroscédasticité pour $\lambda_u = 0$, c'est-à-dire

1) $f(X_t) = \exp(\Sigma_m \delta_m X_{m_t}^{(\lambda_{em})})$,

nous estimons également une forme particulière de celle-ci, à savoir la forme classique de l'hétéroscédasticité, c'est-à-dire

2) $f(X_t) = \Pi_m X_{m_t}^{\delta_m}$, avec $\delta_m = 2$.

3.2. Nomenclature des hypothèses.

Dans la présente sous-section, nous allons formuler différentes hypo-

thèses sur les parties fixe et aléatoire du modèle général, lesquelles vont nous permettre de juger de l'utilité de celui-ci sur une base pragmatique. L'application de ces hypothèses va générer des résultats qui vont nous permettre de faire la distinction entre les deux forces qui affectent la forme fonctionnelle à savoir l'hétéroscédasticité de l'erreur et la non-linéarité. D'une façon plus méthodique, nous pouvons dire que la spécification générale du modèle permet de faire un discernement entre

- dans quelle mesure les paramètres estimés de la forme fonctionnelle compensent pour l'hétéroscédasticité et,
- dans quelle mesure ils reflètent la non-linéarité réellement présente dans le modèle (Gaudry, 1978, p.17).

De plus, l'application des transformations paramétriques Box-Cox à la partie systématique et à la partie stochastique du modèle va induire des résultats qui vont nous instruire sur les effets de la forme fonctionnelle "optimale".

3.2.1. Hypothèses sur la partie systématique du modèle.

En ce qui concerne la partie systématique du modèle, nous établissons six séries d'hypothèses différentes. Dans la première série contrairement aux cinq autres, les paramètres de puissance sont fixés.

Série *: Dans cette série, nous postulons que la transformation Box-Cox sur la variable dépendante, λ_y et la transformation Box-Cox sur chaque groupe de variables indépendantes, λ_{xh} est fixée et contrainte à être égale à zéro ou à l'unité.

Série A: Dans la série A, nous faisons également une restriction d'égalité entre tous les éléments du vecteur λ, sauf que nous ne contraignons plus ces derniers à emprunter la forme linéaire ou log-linéaire. En

laissant les paramètres contraints entre eux s'estimer librement, nous obtiendrons les valeurs optimales de ces paramètres.

Série B: A dessein de distinguer le comportement de la transformation Box-Cox sur la variable dépendante et certaines variables indépendantes du modèle, nous allons libérer ici la contrainte d'égalité entre λ_y et les paramètres λ_x $(\lambda_{x(1)}, \ldots, \lambda_{x(5)})$; ces derniers sont cependant contraints à être égaux entre eux.

Afin de distinguer méthodiquement le comportement des divers groupes de variables exogènes, nous allons décomposer ceux-ci en différents ensembles de combinaisons.

Série C: La série C suggère le relâchement de la contrainte d'égalité entre le groupe des variables de niveau d'activité et les groupes prix, revenu, temps et autres.

Série D: La série D, quant à elle, identifie trois groupes de λ_x. Dans le premier groupe, nous intégrons les variables de prix, de revenu et de temps, tandis que dans les deux autres nous intégrons respectivement les variables d'activité et les attributs divers.

Série E: Cette série est nettement la moins restrictive, car nous n'imposons aucune contrainte à priori sur les lambdas x. Chacun des paramètres varie indépendamment des autres.

3.2.2. Hypothèses sur la partie stochastique du modèle.

A chacune des hypothèses émises sur la partie fixe, nous allons associer six hypothèses différentes sur la partie aléatoire. Ces hypothèses se définissent comme suit:

Hypothèse 2: BC_S

Le symbole BC se réfère à la transformation Box-Cox émise sur la partie fixe du modèle et l'indice S représente les différentes séries énumérées précédemment[4]. Ainsi S = *, A, B, C, D, E, **, AA, BB, CC, DD, EE.

Cette hypothèse suppose l'existence d'une contrainte d'égalité entre chacun des paramètres (ρ_ℓ, λ_{em}, δ_m) de la partie non-systématique du modèle. Cette contrainte conditionne tous les paramètres à égaler zéro.

Hypothèse 3: $BC_S AU$

La même restriction qu'en 2 sauf que nous libérons la contrainte ($\rho_\ell = (\rho_1, \rho_3, \rho_4, \rho_{12}) = 0$) sur la structure d'autocorrélation des erreurs résiduelles (AU). $\rho_1, \rho_3, \rho_4, \rho_{12}$ sont alors estimés en étant non contraints entre eux.

Hypothèse 4: $BC_S HG_\xi$

Cette hypothèse définit un modèle dans lequel nous postulons la formulation générale de l'hétéroscédasticité (HG_ξ) où le symbole ξ

= 1 si la variable $X_{13} \equiv$ *SHONF* est choisie comme source d'hétéroscédasticité

= 2 si la variable $X_{17} \equiv$ *TEND* est choisie comme source d'hétéroscédasticité

= 1,2 si les deux variables sont considérées conjointement comme génératrices de l'hétéroscédasticité.

Les paramètres associés au(x) facteur(s) d'hétéroscédasticité considéré(s) sont alors estimés.

Hypothèse 5: $BC_S HC_\xi$

Ce modèle, qui est un cas spécial de 4, introduit la présence de l'hétéroscédasticité classique (HC_ξ). Comme cette formulation de l'hétéroscédasticité postule l'existence de restrictions contraignantes sur les paramètres de la partie aléatoire, seuls les paramètres de la partie fixe sont estimés.

4.- Si les deux variables d'hétéroscédasticité sont considérées dans $f(X_t)$, alors la série * devient la série **, A devient la série AA et ainsi de suite.

Hypothèse 6: $BC_S AUHG_\xi$

Ce modèle est le moins conditionnel pour chacune des séries car il généralise la procédure Box-Cox aux erreurs qui suivent un schéma autorégressif.

Hypothèse 7: $BC_S AUHC_\xi$

L'hypothèse 7 est similaire à l'hypothèse 5 sauf qu'elle tient compte de l'autocorrélation des erreurs résiduelles.

Pour faciliter la compréhension des différents tests à effectuer, nous mettons en relief dans le tableau 2 la sélection des critères utilisés.

TABLEAU 2

Tableau récapitulatif des diverses hypothèses émises.

Test	Modèle	Hypothèses sur la partie systématique	Hypothèses sur la partie aléatoire
2*	BC_*	$\lambda_y=\lambda_{x1}=\lambda_{x2}=\lambda_{x3}=\lambda_{x4}=\lambda_{x5}$ = 0 ou = 1	$\rho_1=\rho_3=\rho_4=\rho_{12}=0;\lambda_{e1}=0,\lambda_{e2}=0;\delta_1=0,\delta_2=0$
3*	BC_*AU		$\hat{\rho}_1,\hat{\rho}_3,\hat{\rho}_4,\hat{\rho}_{12}$; $\lambda_{e1}=0,\lambda_{e2}=0;\delta_1=0,\delta_2=0$
4*	BC_*HG_1		$\rho_1=\rho_3=\rho_4=\rho_{12}=0;\hat{\lambda}_{e1}$, $\lambda_{e2}=0;\hat{\delta}_1$, $\delta_2=0$
	BC_*HG_2		$\rho_1=\rho_3=\rho_4=\rho_{12}=0;\lambda_{e1}=0,\hat{\lambda}_{e2}$; $\delta_1=0,\hat{\delta}_2$
5*	BC_*HC_1		$\rho_1=\rho_3=\rho_4=\rho_{12}=0;\lambda_{e1}=0,\lambda_{e2}=0;\delta_1=2,\delta_2=0$
	BC_*HC_2		$\rho_1=\rho_3=\rho_4=\rho_{12}=0;\lambda_{e1}=0,\lambda_{e2}=0;\delta_1=0,\delta_2=2$
6*	BC_*AUHG_1		$\hat{\rho}_1,\hat{\rho}_3,\hat{\rho}_4,\hat{\rho}_{12}$; $\hat{\lambda}_{e1}$, $\lambda_{e2}=0;\hat{\delta}_1$, $\delta_2=0$
	BC_*AUHG_2		$\hat{\rho}_1,\hat{\rho}_3,\hat{\rho}_4,\hat{\rho}_{12}$; $\lambda_{e1}=0,\hat{\lambda}_{e2}$; $\delta_1=0,\hat{\delta}_2$
7*	BC_*AUHC_1		$\hat{\rho}_1,\hat{\rho}_3,\hat{\rho}_4,\hat{\rho}_{12}$; $\lambda_{e1}=0,\lambda_{e2}=0;\delta_1=2,\delta_2=0$
	BC_*AUHC_2		$\hat{\rho}_1,\hat{\rho}_3,\hat{\rho}_4,\hat{\rho}_{12}$; $\lambda_{e1}=0,\lambda_{e2}=0;\delta_1=0,\delta_2=2$

2A ⋮ 7A	BC_A $BC_A AUHC_2$	$\hat{\lambda}_y=\hat{\lambda}_{x1}=\hat{\lambda}_{x2}=\hat{\lambda}_{x3}=\hat{\lambda}_{x4}=\hat{\lambda}_{x5}$	mêmes hypothèses mêmes hypothèses
2B ⋮ 7B	BC_B $BC_B AUHC_2$	$\hat{\lambda}_y\neq\hat{\lambda}_{x1}=\hat{\lambda}_{x2}=\hat{\lambda}_{x3}=\hat{\lambda}_{x4}=\hat{\lambda}_{x5}$	mêmes hypothèses mêmes hypothèses
2C ⋮ 7C	BC_C $BC_C AUHC_2$	$\hat{\lambda}_y\neq\hat{\lambda}_{x1}=\hat{\lambda}_{x2}=\hat{\lambda}_{x3}=\hat{\lambda}_{x5}\neq\hat{\lambda}_{x4}$	mêmes hypothèses mêmes hypothèses
2D ⋮ 7D	BC_D $BC_D AUHC_2$	$\hat{\lambda}_y\neq\hat{\lambda}_{x1}=\hat{\lambda}_{x2}=\hat{\lambda}_{x3}\neq\hat{\lambda}_{x4}\neq\hat{\lambda}_{x5}$	mêmes hypothèses mêmes hypothèses
2E ⋮ 7E	BC_E $BC_E AUHC_2$	$\hat{\lambda}_y\neq\hat{\lambda}_{x1}\neq\hat{\lambda}_{x2}\neq\hat{\lambda}_{x3}\neq\hat{\lambda}_{x4}\neq\hat{\lambda}_{x5}$	mêmes hypothèses mêmes hypothèses
4**	$BC_{**}HG_{1,2}$	mêmes hypothèses que *	$\rho_1=\rho_3=\rho_4=\rho_{12}=0;\hat{\lambda}_{e1}\ ,\hat{\lambda}_{e2}\ ;\hat{\delta}_1\ ,\hat{\delta}_2$
5**	$BC_{**}HC_{1,2}$		$\rho_1=\rho_3=\rho_4=\rho_{12}=0;\lambda_{e1}=0,\lambda_{e2}=0;\delta_1=2,\delta_2=2$
6**	$BC_{**}AUHG_{1,2}$		$\hat{\rho}_1,\hat{\rho}_3,\hat{\rho}_4,\hat{\rho}_{12}\ ;\hat{\lambda}_{e1}\ ,\hat{\lambda}_{e2}\ ;\hat{\delta}_1\ ,\hat{\delta}_2$
7**	$BC_{**}AUHC_{1,2}$		$\hat{\rho}_1,\hat{\rho}_3,\hat{\rho}_4,\hat{\rho}_{12}\ ;\lambda_{e1}=0,\lambda_{e2}=0;\delta_1=2,\delta_2=2$
AA		mêmes hypothèses que A	mêmes hypothèses que **
BB		mêmes hypothèses que B	
CC		mêmes hypothèses que C	
DD		mêmes hypothèses que D	
EE		mêmes hypothèses que E	

3.3. Comparaison des hypothèses.

Par un choix judicieux des différentes hypothèses émises sur les parties systématique et stochastique du modèle, il est possible de comparer les valeurs obtenues des maxima de vraisemblance deux à deux avec l'aide du test du rapport des vraisemblances.

Les différentes combinaisons d'hypothèses les plus susceptibles de révéler un certain intérêt sont les suivantes:

. en absence d'autocorrélation.

T. 1 $BC_{\star}$ VS $BC_{\star}HG_1$

(ce test peut être fait pour - $BC_{\star}$ VS $BC_{\star}HG_2$

- $BC_{\star}$ VS $BC_{\star\star}HG_{1,2}$)

$$[-\hat{L}(\lambda_y=\lambda_{x1}=\lambda_{x2}=\lambda_{x3}=\lambda_{x4}=\lambda_{x5}=0 \text{ (ou=1)};\rho_1=\rho_3=\rho_4=\rho_{12}=0;\lambda_{e1}=0,\lambda_{e2}=0;\delta_1=0,\delta_2=0)$$
$$+\hat{L}(\lambda_y=\lambda_{x1}=\lambda_{x2}=\lambda_{x3}=\lambda_{x4}=\lambda_{x5}=0 \text{ (ou=1)};\rho_1=\rho_3=\rho_4=\rho_{12}=0;\hat{\lambda}_{e1},\lambda_{e2}=0;\hat{\delta}_1,\delta_2=0)]$$

T. 2 $BC_{\star}$ VS BC_A [5]

T. 3 BC_A VS $BC_{\star}HG_1$

(ce test peut être fait pour - BC_A VS $BC_{\star}HG_2$

- BC_A VS $BC_{\star\star}HG_{1,2}$)

T. 4 BC_A VS BC_AHG_1

(ce test peut être fait pour - BC_A VS BC_AHG_2

- BC_A VS $BC_{AA}HG_{1,2}$)

T. 5 BC_A VS BC_AHC_1

5.- Pour le test 2 (T.2) et les suivants, voir la liste des hypothèses dans le tableau 2.

(ce test peut être fait pour - BC_A VS BC_AHC_2

- BC_A VS $BC_{AA}HC_{1,2}$)

T. 6 BC_AHC_1 VS BC_AHG_1

(ce test peut être fait pour - BC_AHC_2 VS BC_AHG_2

- $BC_{AA}HC_{1,2}$ VS $BC_{AA}HG_{1,2}$)

T. 7 BC_AHG_1 VS $BC_{AA}HG_{1,2}$

(ce test peut être fait pour - BC_AHG_2 VS $BC_{AA}HG_{1,2}$)

T. 8 BC_AHC_1 VS $BC_{AA}HC_{1,2}$

(ce test peut être fait pour - BC_AHC_2 VS $BC_{AA}HC_{1,2}$)

. en présence d'autocorrélation.

T. 9a BC_* VS BC_*AU

T. 9b BC_A VS BC_AAU

T.10 BC_*AU VS BC_*AUHG_1

(ce test peut être fait pour - BC_*AU VS BC_*AUHG_2

- BC_*AU VS $BC_{**}AUHG_{1,2}$)

T.11 BC_*AU VS BC_AAU

T.12 BC_AAU VS BC_*AUHG_1

(ce test peut être fait pour - BC_AAU VS BC_*AUHG_2

- BC_AAU VS $BC_{**}AUHG_{1,2}$)

T.13 BC_AAU VS BC_AAUHG_1

(ce test peut être fait pour - BC_AAU VS BC_AAUHG_2

- BC_AAU VS $BC_{AA}AUHG_{1,2}$)

T.14 BC_AAU VS BC_AAUHC_1

(ce test peut être fait pour - $BC_A AU$ VS $BC_A AUHC_2$

- $BC_A AU$ VS $BC_{AA} AUHC_{1,2}$)

T.15 $BC_A AUHC_1$ VS $BC_A AUHG_1$

(ce test peut être fait pour - $BC_A AUHC_2$ VS $BC_A AUHG_2$

- $BC_{AA} AUHC_{1,2}$ VS $BC_{AA} AUHG_{1,2}$)

T.16 $BC_A AUHG_1$ VS $BC_{AA} AUHG_{1,2}$

(ce test peut être fait pour - $BC_A AUHG_2$ VS $BC_{AA} AUHG_{1,2}$)

T.17 $BC_A AUHC_1$ VS $BC_{AA} AUHC_{1,2}$

(ce test peut être fait pour - $BC_A AUHC_2$ VS $BC_{AA} AUHC_{1,2}$).

Les tests T.2, T.3, T.4, T.5, T.6, T.7, T.8, T.9b, T.11, T.12, T.13, T.14, T.15, T.16 et T.17 peuvent être faits pour les autres séries.

Les tests T.1 et T.10 servent à détecter la présence de l'hétéroscédasticité ou de l'homoscédasticité des résidus lorsque les paramètres de la partie fixe sont contraints à une valeur particulière.

Les tests T.2 et T.11 permettent de vérifier si la forme fonctionnelle obtenue en laissant les paramètres s'estimer librement est statistiquement différente de la forme fonctionnelle de la situation linéaire ou logarithmique.

Les tests T.3 et T.12 révèlent l'influence de la forme fonctionnelle de la partie fixe comme outil de compensation pour l'hétéroscédasticité.

Les tests T.4 et T.13 confèrent l'utilité de la détermination simultanée de la forme fonctionnelle des parties fixe et aléatoire.

Les tests T.5 et T.14 trouvent leur seule justification en ce qu'ils nous permettent de constater si la valeur de la fonction de vraisemblance croît ou décroît lorsqu'une forme particulière d'hétéroscédasticité est imposée et

si la forme fonctionnelle de la partie systématique varie d'une manière sensible.

Les tests T.6 et T.15 permettent de discerner l'importance de bien déterminer la vraie forme de l'hétéroscédasticité, de révéler l'influence des contraintes sur λ_{em} (=0) et sur δ_m (=2).

Les tests T.7 et T.16 de même que les tests T.8 et T.17 vont nous permettre de voir si la prise en considération simultanée de deux variables comme sources d'hétéroscédasticité produit un impact plus significatif sur la fonction de vraisemblance que la prise en considération d'une seule variable dans $f(X_t)$.

Le test T.9a est effectué dans le but de mesurer le degré d'autocorrélation des erreurs résiduelles dans la fonction, de révéler l'influence de la contrainte sur les $\rho_\ell = \rho_1 = \rho_3 = \rho_4 = \rho_{12}$; le test T.9b sert, en plus, à jauger l'impact de la forme fonctionnelle sur l'autocorrélation.

Jusqu'à présent, l'approche considérée pour établir la comparaison des modèles reposait sur les hypothèses (conditionnelles et inconditionnelles) inhérentes au terme résiduel. En seconde analyse, nous croyons qu'il importe de préciser une autre approche qui va nous permettre de porter un jugement explicite sur la nécessité ou la non-nécessité de libérer les contraintes sur la partie fixe du modèle. Cependant, pour établir des comparaisons entre les séries A et B, B et C et les autres, il devient impératif de ne pas modifier les restrictions sur les paramètres de la variable aléatoire (c'est-à-dire comparer BC_AHG_1 avec BC_BHG_1, etc.).

La statistique pertinente qui va nous permettre de vérifier si l'impact sur la fonction de vraisemblance est significatif (ou marginal) pour chacun des tests énumérés ci-dessus est:

$$-(L_0^{***} - L_1^{***}) > \frac{1}{2}\chi^2_\alpha \text{ (D.L.)}. \qquad (52)$$

Pour le processus de maximisation et pour un test particulier, L_0^{***} représente la valeur logarithmique du maximum de la fonction de vraisemblance associée au modèle le plus contraint et L_1^{***} représente la valeur logarithmique du maximum de la fonction de vraisemblance associée au modèle le moins contraint; χ^2_α (D.L.) dénote la distribution asymptotique chi-carré pour certains degrés de liberté et un niveau de signification α.

Le nombre de degrés de liberté est égal à la différence dans le nombre de paramètres estimés impliqués dans les deux modèles[6].

Les points de pourcentage d'une distribution $\chi^2/2$ pour certains degrés de liberté sont:

Niveau / d.l.	0.10	0.05	0.025	0.01	0.005	0.001	0.0005	0.0001
1	1.36	1.92	2.51	3.32	3.94	5.41	6.05	7.57
2	2.31	3.00	3.69	4.61	5.30	6.91	7.60	9.21
3	3.13	3.91	4.68	5.67	6.42	8.15	8.85	
4	3.89	4.75	5.57	6.64	7.43	9.25	10.00	11.76
5	4.62	5.54	6.42	7.55	8.38	10.25	11.05	

6.- Cependant, certains tests ne respectent pas ce critère.
. Ainsi les tests T.1, T.4, T.10, T.13 (lorsqu'une seule variable est impliquée dans $f(X_t)$) par exemple, requièrent l'utilisation d'un seul degré de liberté même si la différence dans le nombre de paramètres estimés est égale à deux.
Considérons le test T.1.
Pour BC_*HG_1, nous avons $\hat{\delta}_1$, $\hat{\lambda}_{e1}$.
Pour BC_* , nous avons $\delta_1=0$, alors λ_{e1} peut prendre n'importe quelle valeur entre -20 et 20. (Dans notre cas, nous le fixons égal à zéro pour l'empêcher d'intervenir dans le processus de minimisation.)

Puisque pour BC_*, au point $\delta_1=0$ la dimension λ_{el} n'existe pas, alors le test se fait seulement sur le delta.

. Les tests T.5, T.8, T.14 et T.17 peuvent être également considérés comme des cas particuliers.

Considérons le test T.5.

Pour BC_AHC_1, nous avons $\delta_1=2$, $\lambda_{el}=0$.

Pour BC_A , nous avons $\delta_1=0$, λ_{el} peut prendre n'importe quelle valeur.

En conséquence, puisque λ_{el} peut prendre n'importe quelle direction et atteindre n'importe quelle valeur lorsque $\delta_1=0$ et puisque nous n'avons pas des hypothèses emboîtées (nested hypothesis) l'une dans l'autre pour le delta, nous avons alors une absence de degrés de liberté et le critère de base pour appliquer le test de la χ^2 n'est pas rencontré.

Aussi, puisqu'à notre connaissance il n'existe pas de test statistique rigoureux qui puisse permettre une certaine comparaison entre les deux modèles, nous pouvons seulement constater si la valeur de la fonction de vraisemblance augmente, diminue ou encore demeure inchangée.

. Les tests T.3 et T.12 présentent un autre problème.

Soit le test T.3.

Pour BC_A , nous avons $\hat{\lambda}_y=\hat{\lambda}_{x1}=\hat{\lambda}_{x2}=\hat{\lambda}_{x3}=\hat{\lambda}_{x4}=\hat{\lambda}_{x5}$; $\delta_1=0$, λ_{el} (n'importe quelle valeur).

Pour BC_*HG_1, nous avons $\lambda_y=\lambda_{x1}=\lambda_{x2}=\lambda_{x3}=\lambda_{x4}=\lambda_{x5}=1$; $\hat{\delta}_1$, $\hat{\lambda}_{el}$.

Pour ce test, nous avons encore une fois, zéro degré de liberté, puisqu'il n'existe aucune différence dans le nombre de paramètres estimés.

Cependant, lorsque nous considérons les séries B, C, D et E à la place de la série A pour le modèle BC_S dans T.3, alors le problème n'existe plus.

CHAPITRE IV

ANALYSE DES RESULTATS.

Le présent chapitre est consacré à une étude exhaustive des résultats économétriques.

4.1. Justification de l'opérationnalité du modèle.

La première étape consiste à préciser si et dans quelle mesure l'estimation de la détermination conjointe de la forme fonctionnelle d'une fonction économique spécifique et la spécification de la partie stochastique des modèles de régression par le programme DSDM est réalisable sur une base opérationnelle pragmatique.

L'évidence empirique suggère que la fonction de demande de transport en commun mise en relief précédemment, répond d'une façon assez exceptionnelle aux critères d'opérationnalité puisqu'à partir de celle-ci, nous obtenons des résultats qui déterminent les valeurs optimales des différents paramètres et qui génèrent la valeur correspondante du maximum de vraisemblance dans approximativement 70% des expériences tentées.

Cependant, il existe des situations pour lesquelles les solutions dérivées ne nous permettent pas de justifier l'utilité rationnelle des instruments intrinsèques au modèle à cause de divers problèmes d'estimation (numériques ou autres).

Le problème le plus souvent rencontré est celui où l'inversion de la matrice de covariance du gradient est impossible à réaliser; à ce jour, la véritable raison d'être du problème demeure inconnue.

Lorsque cette situation se présente, nous pouvons obtenir la valeur des différents paramètres[1] et la valeur finale du maximum de vraisemblance, mais nous ignorons alors les valeurs des élasticités et des students des variables explicatives. Il est intéressant de noter que les inversions non réussies surviennent surtout lorsque les lambdas de la partie fixe sont tous contraints entre eux et fixés à une certaine valeur et lorsque les contraintes sur les groupes de variables explicatives sont libérées en totalité.

Il existe également des situations où la robustesse du processus par rapport aux modifications de la solution initiale ne généralise pas le concept de convergence numérique vers une solution unique. Ces situations prévalent lorsque les modifications des valeurs initiales pour les paramètres estimés déterminent des valeurs différentes pour les paramètres et paradoxalement génèrent des valeurs du maximum de vraisemblance différentes. Nous avons rencontré ce problème avec le modèle BC_AHG_1 où, pour des valeurs initiales différentes des lambdas ($\lambda(1)$ et $\lambda(2)$), nous avons retrouvé des valeurs divergentes pour les fonctions de vraisemblance, soit respectivement de 273.25 (inversion non réussie) et de 275.83.

La nature d'un troisième problème est axée sur la non-convergence des paramètres. Cette situation survient lorsqu'un (ou des) paramètre(s) introduit(s) dans le processus de maximisation atteint (atteignent) une des bornes de la région admissible. Les modèles BC_CHG_1 et BC_CHC_1 de même que le modèle

---

1.- Il importe de préciser ici que durant la procédure de maximisation pour L^{***}, une contrainte sur chaque paramètre ρ_ℓ est introduite implicitement par un changement de l'espace pour ρ_ℓ. Chacun des ρ_ℓ est alors exprimé en termes de la transformation z de Fisher où,

$$z_\ell = \frac{1}{2} \ln \frac{1 + \rho_\ell}{1 - \rho_\ell} \quad ,$$

avec $-1 < \rho_\ell < 1$ et $-\infty < z_\ell < +\infty$.

$BC_{**}HG_{1,2}$ sont révélateurs de ce problème puisque le paramètre sur le deuxième groupe de variables explicatives, à savoir le revenu, atteint la borne supérieure ($\hat{\lambda}_{x(2)} = 20$) dans les deux premiers cas et le paramètre de puissance sur la variable aléatoire atteint la borne inférieure ($\hat{\lambda}_{e2} = -20$) dans le dernier cas.

Le dernier problème rencontré est celui de la non-respectabilité de la propriété qui veut que la valeur de la fonction de vraisemblance s'accroisse lorsque nous augmentons la généralité de la famille paramétrique dans la procédure de maximisation. Nous faisons face à cet épineux problème lorsque nous estimons le modèle BC_SHG_1 où S = A, B et D. Ainsi pour BC_AHG_1, la valeur finale de la fonction de vraisemblance s'élève à 275.83, pour BC_BHG_1 à 264.74 et pour BC_DHG_1 à 259.45.

Nonobstant la présence de ces quelques problèmes, nous pouvons affirmer en général, que le processus est assez puissant pour générer une solution correspondant à un maximum (au moins local) de la fonction de vraisemblance et que cette fonction de vraisemblance se comporte assez bien dans l'espace des paramètres. En effet, compte tenu du nombre d'observations (169) et du nombre de paramètres relativement élevé impliqués dans la procédure de maximisation, nous pouvons noter que la vitesse de résolution des problèmes est assez rapide.

Le tableau ci-dessous donne un bref aperçu du nombre d'itérations, du temps d'exécution et des coûts monétaires inhérents à certains modèles.

TABLEAU 3

Nombre d'itérations, temps d'exécution et coûts monétaires.

Modèle	Nombre de paramètres	Nombre d'itérations	Temps d'exécution (en secondes)	Coûts monétaires (en dollars)
$BC_A HG_2$	3	22	364.135	40.965
$BC_A AUHG_2$	7	18	556.183	62.570
$BC_D AUHG_1$	9	31	1,001.831	112.705
$BC_E HG_2$	8	26	393.746	44.296
$BC_{AA} HG_{1,2}$	5	45	1,362.082	153.234
$BC_{CC} AUHG_{1,2}$	11	25	1,311.834	147.581

4.2. Estimation et tests de la forme fonctionnelle.

La seconde étape de ce chapitre consiste à donner une réponse dogmatique au problème posé dans la partie théorique. Nous verrons alors l'utilité de tester la forme fonctionnelle de la partie fixe simultanément avec celle de la partie stochastique.

4.2.1. Justification pratique de l'utilité de la détermination conjointe par une comparaison des modèles à partir des hypothèses émises sur la partie aléatoire.

Afin de discerner formellement l'utilité de la détermination conjointe sur cette base, nous allons montrer à partir des résultats de régression les plus significatifs, de quelle manière la forme fonctionnelle peut s'avérer fausse si le lambda de la partie systématique du modèle exerce un rôle de compensation pour l'hétéroscédasticité. Nous verrons alors l'utilité de la transformation Box-Cox sur la partie fixe simultanément à la paramétrisation de la partie stochastique pour la détermination de la forme fonctionnelle optimale. (Les valeurs des paramètres ρ_ℓ, λ_{em}, δ_m qui spécifient la distribution des ré-

sidus ne sont pas indépendantes de celles qui gouvernent la forme fonctionnelle, elles doivent donc être examinées conjointement.)

Les tests requis pour nous permettre de dégager des conclusions pratiques à une théorie qui semble implicite dans ses fondements sont les tests T.1, T.2, T.3 et T.4 lorsque nous ignorons l'autocorrélation des erreurs résiduelles et les tests T.9a, T.10, T.11, T.12 et T.13 lorsque nous considérons le schéma autorégressif des erreurs résiduelles.

Tout d'abord, nous allons voir quel est l'impact sur la valeur de la fonction de vraisemblance et sur la forme fonctionnelle lorsque nous considérons la variable *SHONF* comme source possible de l'hétéroscédasticité de l'erreur.

Le tableau 4 présente les valeurs des paramètres et les valeurs associées des fonctions de vraisemblance.

Considérons le cas suivant:

CAS I

Les résultats présentés dans le tableau 4 montrent que le principal impact sur la valeur de la fonction de vraisemblance provient de la présence significative de l'autocorrélation des erreurs résiduelles.

En effet, la valeur du test statistique pour la comparaison des modèles $BC_\star$ et $BC_\star AU$ (test T.9a) s'élève à 29.72, ce qui signifie que nous pouvons rejeter le modèle qui ne tient pas compte de l'autocorrélation des résidus avec un très haut degré de confiance.

Cette première constatation s'avère réaliste si l'on tient compte du fait qu'en séries chronologiques, l'on doit s'attendre à des liaisons tempo-

TABLEAU 4

Valeurs de λ_y, λ_{xk}, ρ_ℓ, λ_{em} et δ_m pour $X_m \equiv$ *SHONF* et valeurs correspondantes des fonctions de vraisemblance. Séries ⋆ et A.

Paramètres	Modèles I $BC_\star$	II $BC_\star AU$	III $BC_\star AUHG_1$	IV $BC_A AU$	V $BC_A AUHG_1$	VI $BC_A AUHC_1$
λ_y	1.00	1.00	1.00	.72 (13.04)[b] (-4.99)[c]	.75 (3.55) (-1.17)	.93 (16.61) (-1.18)
λ_x[a]	1.00	1.00	1.00	.72	.75	.93
ρ_1	0.00	.17 (2.26)	n.d.	.14 (1.82)[b]	.11 (1.45)	.05 (.67)
ρ_3	0.00	-.32 (-3.64)	n.d.	-.34 (-3.73)	-.33 (-3.36)	-.28 (-2.67)
ρ_4	0.00	-.07 (-.87)	n.d.	-.09 (-1.08)	-.10 (-1.20)	-.14 (-1.60)
ρ_{12}	0.00	.41 (5.19)	n.d.	.38 (4.66)	.39 (4.67)	.38 (4.81)
λ_{e1}	0.00	0.00	n.d.	0.00	-.46 (- .03) (- .10)	0.00
λ_{e2}	0.00	0.00	n.d.	0.00	0.00	0.00
δ_1	0.00	0.00	n.d.	0.00	.17 (.11)	2.00
δ_2	0.00	0.00	n.d.	0.00	0.00	0.00
L***	237.10	266.82	n.d.	268.00	268.07	264.00

a) Pour $\lambda_{x(1)}$, $\lambda_{x(2)}$, $\lambda_{x(3)}$, $\lambda_{x(4)}$, $\lambda_{x(5)}$.
b) Le test-t calculé sous l'hypothèse que le paramètre est égal à zéro.
c) Le test-t calculé sous l'hypothèse que le paramètre est égal à un. Ce test est calculé pour λ_y, λ_x, λ_{em}.

UNIVERSITÉ DE MONTRÉAL

relles entre les erreurs commises sur les variables.

a) Par ailleurs, pour détecter la présence de l'hétéroscédasticité pour le modèle linéaire, nous devons appliquer le test T.10 dans lequel interfèrent les modèles BC_*AU et BC_*AUHG_1. Cependant, à cause du problème d'inversion de la matrice de covariance du gradient, nous ne possédons pas la valeur exacte du logarithme de la fonction de vraisemblance concentrée pour ce dernier modèle. Néanmoins, considérant que la valeur de L^{***} pour le modèle BC_*AUHG_1 se doit d'être inférieure à celle du modèle BC_AAUHG_1 à cause d'une différence dans le nombre de paramètres estimés, nous pouvons prendre pour acquis que la valeur du test statistique est inférieure à 1.25.

Puisque cette valeur est inférieure à 1.36, soit la valeur de la région critique pour un niveau de confiance de 90%, nous pouvons rejeter la présence de l'hétéroscédasticité pour le modèle linéaire.

b) Nous allons étudier ici l'effet sur la forme fonctionnelle lorsque nous laissons estimer librement les paramètres de la partie fixe.

La colonne IV du tableau 4 montre que la libération de la contrainte de fixité entraîne une diminution dans la valeur de λ_y, celui-ci passant d'une valeur de 1 à une valeur de .72. La différence entre les modèles BC_*AU et BC_AAU n'étant pas significative (1.18) et la valeur des coefficients autorégressifs demeurant pratiquement inchangée, le test non-conditionnel sur λ_y ne nous permet pas de rejeter le modèle linéaire à un niveau de confiance d'au moins 90%. De plus, comme $\hat{\lambda}_y$ n'est pas biaisé par un rôle de compensation pour l'hétéroscédasticité, nous pouvons supposer qu'une forme linéaire ne peut être rejetée d'une façon significative.

c) Les hypothèses faites sur le terme d'erreur (à savoir l'homoscédasticité) et sur la forme fonctionnelle (à savoir une forme approximativement li-

néaire) sont soutenues par l'estimation du modèle qui considère simultanément les transformations sur les parties systématique et stochastique. En effet, les résultats obtenus pour ce modèle ($BC_A AUHG_1$) révèlent que les erreurs sont homoscédastiques et que la forme fonctionnelle optimale est approximativement linéaire ($\hat{\lambda}_y = .75$, ce qui n'est pas statistiquement différent de l'unité d'après le test-t de student).

Pour les fins de notre discussion, nous allons maintenant exposer et analyser les résultats qui sont générés lorsque la variable *TEND* est impliquée dans la formulation générale de l'hétéroscédasticité. (Voir tableau 5).

CAS II

a) La comparaison des modèles BC_* et $BC_* HG_2$ nous permet de constater que la variable *TEND* constitue au moins un des éléments (peut-être constitue-t-elle le seul obstacle) qui nous empêche de postuler l'existence d'une variance constante pour le terme résiduel, dans un modèle linéaire.

Le modèle BC_* pouvant être rejeté à un niveau de signification de .0001 par le test du rapport des vraisemblances (12.27 > 7.57 pour 1 degré de liberté), cela témoigne de la pertinence des tests sur les paramètres λ_{em} et δ_m.

b) Comme dans le cas I, nous allons maintenant relâcher la contrainte des λ fixes pour la partie systématique du modèle. En agissant ainsi, nous retrouvons une forme fonctionnelle "racine carrée" puisque la valeur de λ_y estimé s'établit à .51. La valeur obtenue pour $\hat{\lambda}_y$ suggère que le paramètre de puissance généré a une certaine capacité à minimiser l'hétéroscédasticité de la variable aléatoire. Nous pouvons donc affirmer que la forme fonctionnelle est faussée de par son rôle compensateur pour l'hétéroscédasticité. Il importe de préciser toutefois que ce paramètre de puissance n'a pas agi comme un

TABLEAU 5

Valeurs de λ_y, λ_{xk}, ρ_ℓ, λ_{em} et δ_m pour $X_m \equiv$ *TEND* et valeurs correspondantes des fonctions de vraisemblance. Séries ⋆ et A.

Paramètres	Modèles				
	I $BC_\star$	II $BC_\star HG_2$	III BC_A	IV $BC_A HG_2$	V $BC_A HC_2$
λ_y	1.00	1.00	.51 (10.51) (-10.07)	.48 (4.20) (-4.50)	1.05 (19.26) (1.00)
λ_x	1.00	1.00	.51	.48	1.05
ρ_1	0.00	0.00	0.00	0.00	0.00
ρ_3	0.00	0.00	0.00	0.00	0.00
ρ_4	0.00	0.00	0.00	0.00	0.00
ρ_{12}	0.00	0.00	0.00	0.00	0.00
λ_{e1}	0.00	0.00	0.00	0.00	0.00
λ_{e2}	0.00	1.85 (2.72) (1.25)	0.00	2.14 (1.83) (.98)	0.00
δ_1	0.00	0.00	0.00	0.00	0.00
δ_2	0.00	.02 (.71)	0.00	.01 (.38)	2.00
L^{***}	237.10	249.37	241.19	253.72	223.36

UNIVERSITÉ DE MONTRÉAL

instrument de correction suffisant. En effet, la correction "complète" pour l'hétéroscédasticité aurait exigée que la valeur de $\hat{\lambda}_y$ soit telle que la fonction de vraisemblance devienne égale (approximativement) à 249.37. Cependant, l'application du test T.3 pour le modèle BC_A versus le modèle BC_*HG_2 nous permet de constater que la valeur de la fonction de vraisemblance du premier modèle est beaucoup plus petite que celle du second modèle: (241.19 < 249.37).

c) La transformation Box-Cox sur la partie fixe n'ayant pas minimisé l'hétéroscédasticité complètement, nous allons introduire l'instrument spécifique à cette fin (à savoir la transformation paramétrique sur la variable susceptible d'engendrer de l'hétéroscédasticité et impliquée dans $f(X_t)$, c'est-à-dire sur la variable *TEND*). Les résultats montrent que la détermination conjointe de la forme fonctionnelle de la partie fixe et de la forme générale de l'hétéroscédasticité a pour effet de faire augmenter la valeur de la fonction de vraisemblance d'une façon appréciable, soit de 241.19 à 253.72. La valeur statistique pour la comparaison des modèles BC_A et BC_AHG_2 se chiffrant à 12.53, nous pouvons accepter avec un très haut degré de confiance (.0001) le modèle[2] BC_AHG_2.

Il est par ailleurs extrêmement intéressant de noter que la valeur de $\hat{\lambda}_y$ (.48) demeure pratiquement inchangée. Aussi, nous voyons très clairement dans ce cas, que la force qui affecte la forme fonctionnelle n'est pas l'hétéroscédasticité de l'erreur mais bien la non-linéarité réellement présente dans le modèle. De plus, il y a une forte présomption à croire que la forme fonctionnelle est robuste à l'hétéroscédasticité, puisqu'elle n'est pas modifiée qu'il y ait présence ou absence de l'hétéroscédasticité de l'erreur.

2.- Il est intéressant de remarquer que la présence de l'hétéroscédasticité est aussi importante dans le non-linéaire que dans le linéaire même si la valeur de λ_y a diminuée de la moitié.

Nous pouvons conclure pour ce cas que l'utilisation rationnelle des instruments (λ_{em} et δ_m) a eu pour effet de changer le caractère hétéroscédastique des erreurs.

Ce cas est également très révélateur en ce qui concerne l'utilité de bien déterminer la forme de l'hétéroscédasticité. La comparaison des modèles BC_AHC_2 et BC_AHG_2 révélant une différence de 30.36 dans L^{***}, ceci implique, avec 2 degrés de liberté, que le modèle BC_AHC_2 peut être rejeté à un niveau de 99.99%. De plus, une comparaison avec le modèle BC_A nous permet d'affirmer ici, qu'il est préférable de ne pas corriger du tout plutôt que d'essayer de corriger avec la formulation classique de l'hétéroscédasticité.

L'utilisation de cette spécification nous aurait permis de conclure à la linéarité du modèle, alors que celui-ci est vraiment non-linéaire.

Le tableau 6 présente les résultats induits lorsqu'il y a la relaxation de la contrainte sur les ρ_ℓ.

Soit le cas suivant:

CAS III

Le cas III démontre une forte analogie avec le cas II, sauf que pour le cas III, la forme fonctionnelle semble particulièrement robuste à la linéarité.

Une comparaison des valeurs du maximum de vraisemblance des modèles BC_*AU et BC_*AUHG_2 nous permet de rejeter avec confiance le modèle linéaire qui n'introduit pas un élément de correction pour l'hétéroscédasticité. L'écart considérable entre les valeurs de L^{***} pour les deux modèles précités, de même que les valeurs générées pour $\hat{\lambda}_{e2}$ (1.08) et $\hat{\delta}_2$ (.12) nous instruisent sur l'importance réelle de corriger pour l'hétéroscédasticité.

TABLEAU 6

Valeurs de λ_y, λ_{xk}, ρ_ℓ, λ_{em} et δ_m pour $X_m \equiv$ *TEND* et valeurs correspondantes des fonctions de vraisemblance. Séries * et A.

Paramètres	Modèles				
	I BC_*AU	II BC_*AUHG_2	III BC_AAU	IV BC_AAUHG_2	V BC_AAUHC_2
λ_y	1.00	1.00	.72 (13.04) (-4.99)	1.004 (5.94) (.03)	1.14 (17.94) (2.21)
λ_x	1.00	1.00	.72	1.004	1.14
ρ_1	.17 (2.26)	.25 (2.92)	.14 (1.82)	.25 (2.94)	-.09 (-.78)
ρ_3	-.32 (-3.64)	-.19 (-1.82)	-.34 (-3.73)	-.19 (-1.77)	-.39 (-5.69)
ρ_4	-.07 (-.87)	-.06 (-.58)	-.09 (-1.08)	-.06 (-.57)	-.27 (-2.53)
ρ_{12}	.41 (5.19)	.47 (6.84)	.38 (4.66)	.47 (6.87)	.15 (2.65)
λ_{e1}	0.00	0.00	0.00	0.00	0.00
λ_{e2}	0.00	1.08 (1.89) (.13)	0.00	1.07 (1.30) (.09)	0.00
δ_1	0.00	0.00	0.00	0.00	0.00
δ_2	0.00	.12 (.98)	0.00	.12 (.60)	2.00
L^{***}	266.82	277.28	268.00	277.28	242.77

UNIVERSITÉ DE MONTRÉAL

En libérant la contrainte de fixité du paramètre lambda, nous obtenons un λ_y qui passe d'une valeur de 1.00 à une valeur de .72. Cependant, comme dans le cas précédent, la valeur du maximum de vraisemblance nous dicte que la transformation Box-Cox ne suffit pas à annihiler l'hétéroscédasticité et qu'il nous faut, par conséquent, rendre plus sophistiquée la spécification du terme résiduel par la prise en considération simultanée de la structure d'autocorrélation et de l'hétéroscédasticité.

A ce moment, il y a une correction complète de l'hétéroscédasticité et la vraie forme du modèle, contrairement au cas II, n'est plus une forme puissance à la 1/2 mais elle est purement linéaire ($\hat{\lambda}_y$ = 1.004), voir colonne IV du tableau 6.

En résumé, les résultats inhérents à ce cas nous permettent d'affirmer que le modèle optimal est vraiment linéaire et de supposer que la forme fonctionnelle est assez robuste à l'hétéroscédasticité.

Jusqu'à présent, nous avons considéré les variables *SHONF* et *TEND* comme étant impliquées isolément dans $f(X_t)$. Nous allons maintenant juger les impacts sur L^{***} et sur λ_y lorsque les deux variables sont choisies conjointement comme sources de l'hétéroscédasticité.

CAS IV

Le tableau 7 fournit les résultats des modèles $BC_{AA}AUHG_{1,2}$ et $BC_{AA}AUHC_{1,2}$.

La prise en considération des deux variables simultanément laisse voir un impact considérable sur la valeur de la fonction de vraisemblance. Ainsi, lorsque les deux variables d'activité sont impliquées conjointement dans $f(X_t)$, la valeur de la fonction de vraisemblance s'élève à 295.77 comparativement à

268.07 et à 277.28 lorsque nous considérons séparément les variables *SHONF* et *TEND* dans la formulation générale.

TABLEAU 7

Valeurs de λ_y, λ_{xk}, ρ_ℓ, λ_{em} et δ_m pour $X_m \equiv$ *SHONF* et *TEND* et valeurs correspondantes des fonctions de vraisemblance. Séries A et AA.

Paramètres	Modèles	
	I $BC_{AA}AUHG_{1,2}$	II $BC_{AA}AUHC_{1,2}$
λ_y	.92 (6.84) (- .61)	1.22 (20.55) (3.66)
λ_x	.92	1.22
ρ_1	.63 (9.24)	.35 (3.18)
ρ_3	-.59 (-8.17)	-.30 (-3.86)
ρ_4	.35 (5.09)	-.02 (-.21)
ρ_{12}	.32 (4.87)	-.09 (-1.81)
λ_{e1}	5.38 (9.10) (7.41)	0.00
λ_{e2}	.36 (.88) (-1.58)	0.00
δ_1	-.03 (-1.43)	2.00
δ_2	.49 (1.56)	2.00
L^{***}	295.77	216.97

Paramètre	BC_AAUHG_1	BC_AAUHG_2
λ_y	.75	1.004
L^{***}	268.05	277.28

BC_AAUHC_1	BC_AAUHC_2
.93	1.14
264.00	242.77

La valeur obtenue pour L^{***} suggère que les variables *SHONF* et *TEND* utilisées conjointement, sont la source d'une très forte présence d'hétéroscédasticité, alors que l'on se rappelle que prise isolément, la variable *SHONF* n'est pas une variable qui cause de l'hétéroscédasticité.

Les valeurs du test du rapport des vraisemblances χ^2 pour la comparaison des modèles $BC_A AUHG_1$ versus $BC_{AA} AUHG_{1,2}$ (test T.16) et $BC_A AUHG_2$ versus $BC_{AA} AUHG_{1,2}$ s'établissent respectivement à 27.72 et à 18.49 ce qui, pour 2 degrés de liberté, nous permet de rejeter les modèles $BC_A AUHG_1$ et $BC_A AUHG_2$ en faveur du modèle $BC_{AA} AUHG_{1,2}$ avec un niveau de confiance extrêmement élevé.

Le tableau 7 montre également que la valeur de L^{***} pour le modèle $BC_{AA} AUHC_{1,2}$ est réduite d'une façon très sensible lorsque nous utilisons la forme multiplicative multivariée classique.

En ce qui a trait à la forme fonctionnelle du modèle $BC_{AA} AUHG_{1,2}$, elle demeure robuste à la linéarité.

4.2.2. Justification de l'utilité de la généralisation de la paramétrisation de la forme fonctionnelle par une comparaison des modèles à partir des hypothèses émises sur la partie systématique.

La prochaine phase de notre exposé consiste à établir une comparaison des différents modèles à partir des modifications opérées sur la partie fixe du modèle. Pour vérifier les impacts sur L^{***} et sur les paramètres λ_y et λ_{xk}, nous limiterons notre analyse aux modèles $BC_S AU$ et $BC_S HG_2$ où S = *, A, B, C, D et E.

Les tableaux 8, 9, 10 et 11 révèlent, à l'intérieur des diverses applications des modèles $BC_S AU$ et $BC_S HG_2$, que les valeurs des fonctions de vraisemblance augmentent au fur et à mesure que nous libérons les contraintes sur les paramètres λ_y et λ_{xk}. Il importe de préciser toutefois que la relaxation

TABLEAU 8

Valeurs de λ_y, λ_{xk}, ρ_ℓ, λ_{em}, δ_m et valeurs correspondantes des fonctions de vraisemblance. - Modèle $BC_S AU$ - Séries *, A, B, C, D, E.

Paramètres	Modèles						
	I $BC_* AU$	II $BC_* AU$	III $BC_A AU$	IV $BC_B AU$	V $BC_C AU$	VI $BC_D AU$	VII $BC_E AU$
λ_y	0.00	1.00	.72 (13.04) (-4.99)	.69 (12.19) (-5.36)	.63 (10.72) (-6.22)	.64 (10.91) (-6.03)	n.d.
$\lambda_{x(1)}$	0.00	1.00	.72	.98 (13.73) (- .03)	5.10 (2.75) (2.21)	5.07 (2.70) (2.17)	n.d.
$\lambda_{x(2)}$	0.00	1.00	.72	.98	5.10	5.07	n.d.
$\lambda_{x(3)}$	0.00	1.00	.72	.98	5.10	5.07	n.d.
$\lambda_{x(4)}$	0.00	1.00	.72	.98	1.20 (3.76) (.63)	1.26 (3.42) (.71)	n.d.
$\lambda_{x(5)}$	0.00	1.00	.72	.98	5.10	.66 (.09) (- .05)	n.d.
ρ_1	.10 (1.33)	.17 (2.26)	.14 (1.82)	.15 (1.94)	.06 (.74)	.08 (1.02)	n.d.
ρ_3	-.34 (-3.88)	-.32 (-3.64)	-.34 (-3.73)	-.33 (-3.53)	-.51 (-6.83)	-.49 (-6.26)	n.d.
ρ_4	-.10 (-1.22)	-.07 (-.87)	-.09 (-1.08)	-.07 (-.91)	-.11 (-1.31)	-.09 (-1.11)	n.d.
ρ_{12}	.33 (4.36)	.41 (5.19)	.38 (4.66)	.40 (4.96)	.28 (4.02)	.29 (3.91)	n.d.
λ_{e1}	0.00	0.00	0.00	0.00	0.00	0.00	
λ_{e2}	0.00	0.00	0.00	0.00	0.00	0.00	
δ_1	0.00	0.00	0.00	0.00	0.00	0.00	
δ_2	0.00	0.00	0.00	0.00	0.00	0.00	
L^{***}	257.36	266.82	268.00	268.15	272.72	272.90	n.d.
Nombre de paramètres	4	4	5	6	7	8	

TABLEAU 9

Valeurs de λ_y, λ_{xk}, ρ_ℓ, λ_{em}, δ_m et valeurs correspondantes des fonctions de vraisemblance. - Modèle BC_SHG_2 - Séries *, A, B, C, D, E.

Paramètres	Modèles						
	I BC_*HG_2	II BC_*HG_2	III BC_AHG_2	IV BC_BHG_2	V BC_CHG_2	VI BC_DHG_2	VII BC_EHG_2
λ_y	0.00	1.00	.48 (4.20) (-4.50)	.44 (3.83) (-4.82)	.42 (3.46) (-4.76)	.37 (2.87) (-4.97)	.32 (2.28) (-4.76)
$\lambda_{x(1)}$	0.00	1.00	.48	.63 (1.08) (- .62)	2.99 (1.31) (.87)	10.35 (2.44) (2.20)	9.05 (4.17) (3.70)
$\lambda_{x(2)}$	0.00	1.00	.48	.63	2.99	10.35	7.07 (.82) (.70)
$\lambda_{x(3)}$	0.00	1.00	.48	.63	2.99	10.35	5.20 (1.53) (1.23)
$\lambda_{x(4)}$	0.00	1.00	.48	.63	.74 (1.73) (- .61)	.21 (.36) (-1.33)	.09 (.33) (-3.25)
$\lambda_{x(5)}$	0.00	1.00	.48	.63	2.99	-.96 (- .20) (- .40)	-.87 (- .19) (- .40)
ρ_1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ρ_3	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ρ_4	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ρ_{12}	0.00	0.00	0.00	0.00	0.00	0.00	0.00
λ_{e1}	0.00	0.00	0.00	0.00	0.00	0.00	0.00
λ_{e2}	2.44 (2.21) (1.31)	1.85 (2.72) (1.25)	2.14 (1.83) (.98)	2.17 (1.78) (.96)	2.02 (1.69) (.85)	1.93 (1.45) (.70)	1.70 (1.23) (.50)
δ_1	0.00	0.00	0.00	0.00	0.00	0.00	0.00
δ_2	.01 (.42)	.02 (.71)	.01 (.38)	.01 (.37)	.02 (.38)	.02 (.34)	.03 (.34)
L***	248.98	249.37	253.72	253.86	255.04	256.09	260.00
Nombre de paramètres	2	2	3	4	5	6	8

TABLEAU 10

Valeurs des tests statistiques χ^2 pour les paramètres λ_y et λ_{xk}. - Modèle BC_SAU - Séries *, A, B, C, D, E.

Modèles						
BC_*AU[a]	BC_*AU[b]	BC_AAU	BC_BAU	BC_CAU	BC_DAU	BC_EAU
0.	9.46 (0)[c]	-	-	-	-	-
	0.	1.18 (1)	-	-	-	-
		0.	.15 (1)	4.72 (2)	4.90 (3)	n.d.[d]
			0.	4.57 (1)	4.75 (2)	n.d.
				0.	.18 (1)	n.d.
					0.	n.d.
						0.

a) Modèle logarithmique.
b) Modèle linéaire.
c) Les chiffres entre parenthèses indiquent les degrés de liberté.
d) Les valeurs pour ces tests ne sont pas disponibles.

TABLEAU 11

Valeurs des tests statistiques χ^2 pour les paramètres λ_y et λ_{xk}. - Modèle BC_SHG_2 - Séries *, A, B, C, D, E.

Modèles						
BC_*HG_2	BC_*HG_2	BC_AHG_2	BC_BHG_2	BC_CHG_2	BC_DHG_2	BC_EHG_2
0.	.39 (0)	-	-	-	-	-
	0.	4.35 (1)	-	-	-	-
		0.	.14 (1)	1.32 (2)	2.37 (3)	6.28 (5)
			0.	1.18 (1)	2.23 (2)	6.14 (4)
				0.	1.05 (1)	4.96 (3)
					0.	3.91 (2)
						0.

des contraintes d'égalité entre les lambdas xk ne provoque pas un accroissement stable et proportionnel de L^{***} relativement au nombre des degrés de liberté. La valeur du maximum de vraisemblance est la plus faible lorsque nous imposons la forme logarithmique ($\lambda_y = \lambda_{x(1)} = \lambda_{x(2)} = \lambda_{x(3)} = \lambda_{x(4)} = \lambda_{x(5)} = 0$). Une amélioration très importante est cependant observée pour le modèle BC_SAU lorsque nous imposons la forme linéaire ($\lambda_y = \lambda_{x(1)} = \lambda_{x(2)} = \lambda_{x(3)} = \lambda_{x(4)} = \lambda_{x(5)} = 1$). La valeur de L^{***} passe alors de 257.36 à 266.82, soit une différence appréciable de 9.46. Par ailleurs, pour ce modèle, si nous conservons l'égalité entre le paramètre λ_y et les paramètres λ_{xk} tout en les laissant se déterminer librement ($\hat{\lambda}_y = \hat{\lambda}_{x(1)} = \hat{\lambda}_{x(2)} = \hat{\lambda}_{x(3)} = \hat{\lambda}_{x(4)} = \hat{\lambda}_{x(5)} = .72$), cela ne produit qu'un accroissement de 1.18 dans la valeur de L^{***}, ce qui pour 1 degré de liberté ne s'avère pas statistiquement significatif à un niveau de confiance d'au moins 0.1. En relâchant les contraintes une à la fois sur les λ_{xk}, nous pouvons remarquer que l'impact le plus important est réalisé lorsque le lambda sur le groupe des variables d'activité ($\lambda_{x(4)}$) est estimé isolément des lambdas affectant les autres groupes de variables explicatives, alors contraints entre eux. L'écart entre $L_0{}^{***}$ et $L_1{}^{***}$ pour les modèles BC_BAU et BC_CAU s'établit à 4.57.

Le modèle BC_SHG_2 ne présente pas un comportement similaire au modèle BC_SAU. Contrairement à ce dernier modèle, le modèle BC_SHG_2 montre un accroissement notable ($\Delta = 4.35$) dans la valeur de la fonction de vraisemblance lorsque nous passons d'une forme linéaire à une forme moins contraignante (c'est-à-dire, une forme où tous les lambdas sont estimés contraints entre eux). Nonobstant, la valeur du test statistique pour la comparaison des différentes applications du modèle BC_SHG_2 dénote une différence encore plus significative lorsque nous généralisons la procédure Box-Cox à tous les λ_{xk} désagrégés. En effet, lorsque la forme la plus générale est générée, nous obtenons une valeur de L^{***} égale à 260.00, soit une amélioration nette de 6.28 (pour 5 degrés de

liberté) par rapport au modèle le plus restrictif sur les λ_{xk} estimés.

Il est intéressant de remarquer, pour les deux modèles, que la libération de la condition d'égalité entre le paramètre lambda y et les paramètres lambdas xk crée l'impact le moins significatif sur la valeur de la fonction de vraisemblance; nous notons des augmentations respectives de .15 et de .14 pour les modèles BC_SAU et BC_SHG_2.

Après avoir étudié les impacts sur L^{***}, nous allons poursuivre nos investigations en examinant brièvement le comportement de la forme fonctionnelle. Considérons le modèle BC_SHG_2.

Trois points peuvent être soulevés.

- Premièrement, les variantes du modèle BC_SHG_2 indiquent que la valeur de $\hat{\lambda}_y$ est assez stable comparativement à celles des $\hat{\lambda}_{xk}$. La faible variation dans la valeur de λ_y reflète probablement la faible variation dans la valeur des fonctions de vraisemblance (pour 1 D.L.). L'explication la plus logique à cette stabilité relative de la fonction de vraisemblance est sans doute reliée à la présence du paramètre λ_y dans le Jacobien.

- Deuxièmement, il semble que le paramètre $\hat{\lambda}_{x(4)}$ soit celui qui est le plus robuste à la désagrégation.

- Troisièmement, les résultats montrent que les valeurs des tests-t de student diminuent, règle générale, au fur et à mesure que nous relâchons les contraintes sur les λ_{xk}. Nous pouvons ajouter, après consultation des tableaux 5, 6 et 7, que les valeurs numériques du test de student liées au paramètre λ_y sont beaucoup plus élevées lorsque nous postulons la formulation classique de l'hétéroscédasticité plutôt que la formulation générale. De plus, les résultats inscrits aux tableaux 4, 5 et 6 semblent indiquer que la correction de l'hétéroscédasticité par la formulation générale fait diminuer les valeurs statistiques du paramètre de puissance λ_y.

4.2.3. Détermination du modèle optimal pour la fonction de demande du transport en commun à Montréal (marché des écoliers).

L'estimation du modèle le plus général, à savoir le modèle[3] $BC_{CC}AUHG_{1,2}$ nous permet de postuler que la forme fonctionnelle qui donne le "meilleur ajustement" pour la fonction de demande de transport en commun à Montréal n'est pas linéaire. En effet, les paramètres qui optimisent la fonction de vraisemblance (296.81) ont atteint les valeurs exposées dans le tableau suivant:

TABLEAU 12

Valeurs des paramètres qui optimisent la fonction de vraisemblance.

Paramètres	$BC_{CC}AUHG_{1,2}$
λ_y	.57 (4.21) (-3.24)
$\lambda_{x(1)}=\lambda_{x(3)}=\lambda_{x(4)}=\lambda_{x(5)}$	.96 (1.51) (- .07)
$\lambda_{x(2)}$	.68 (2.07) (- .97)
ρ_1	.60 (7.94)
ρ_3	-.41 (-4.39)
ρ_4	.11 (1.21)
ρ_{12}	.32 (4.26)
λ_{e1}	4.82 (7.71) (6.11)
λ_{e2}	1.89 (1.72) (.81)
δ_1	-.05 (-1.39)
δ_2	.02 (.42)
L^{***}	296.81

3.- A cause du problème d'inversion de la matrice de covariance du gradient, nous n'avons pu obtenir les estimés du modèle $BC_{EE}AUHG_{1,2}$.

Cependant, comme la valeur statistique pour le test du rapport des vraisemblances ne nous permet pas de rejeter le modèle $BC_{AA}AUHG_{1,2}$ à un niveau de confiance élevé (la valeur pour ce test s'élève à 1.04), nous pouvons tout aussi bien accepter ce dernier modèle qui présente une forme fonctionnelle approximativement doublement linéaire. Nous rappelons que dans le modèle BC_{AA}-$AUHG_{1,2}$, la valeur de $\bar{\lambda}_y$ ($= \bar{\lambda}_{xk}$) s'élève à .92.

4.3. Divers effets de la forme fonctionnelle de la partie systématique et de la spécification de l'erreur.

Si les résultats analysés jusqu'à présent confirment l'utilité d'utiliser la détermination simultanée de la forme fonctionnelle de la partie fixe et la spécification de la formulation générale de l'hétéroscédasticité, nous allons maintenant examiner l'influence de la forme fonctionnelle des parties fixe et aléatoire sur les estimations des paramètres des variables fixes et aléatoires du modèle. A cette fin, nous analyserons les effets de la forme fonctionnelle sur la distribution des résidus autorégressifs, sur la signification des variables, sur les élasticités et les signes des paramètres ainsi que sur les coefficients autorégressifs. Nous verrons alors si une mauvaise spécification de la forme fonctionnelle et du terme de l'erreur peut donner des résultats aberrants et non indicatifs de la réalité ou de nos anticipations.

4.3.1. Effets de la forme fonctionnelle et de la spécification de l'erreur sur la distribution des erreurs résiduelles autorégressives.

Dans cette sous-section, nous limiterons notre analyse aux effets de la forme fonctionnelle de la partie fixe et/ou de la partie aléatoire sur la distribution des erreurs autorégressives, puisque nous avons déjà discuté amplement des impacts de la forme fonctionnelle de la partie fixe sur les résidus hétéroscédastiques (v.g. la correction de l'hétéroscédasticité, par exemple).

La valeur du test statistique pour la comparaison des modèles BC_* et BC_A s'élève à 4.09; ceci implique, avec 1 degré de liberté, que les deux modèles sont statistiquement différents à un niveau de 99.5% et signifie que les ordres du schéma autorégressif postulés pour $\lambda_y = \lambda_{xk} = 1$; $\rho = 0$; $\lambda_e = 0$; $\delta = 0$ peuvent être différents (statistiquement) de ceux qui auraient pu être postulés pour $\hat{\lambda}_y = \hat{\lambda}_{xk} = .51$; $\rho = 0$; $\lambda_e = 0$; $\delta = 0$. Cependant, une analyse des fonctions d'autocorrélation et d'autocorrélation partielle (tableau 13) montre qu'il n'y a aucun changement significatif dans la structure, lorsque nous passons d'un modèle linéaire dans (a) à un modèle non-linéaire dans (c). La force de la structure n'est donc pas fondamentalement affaiblie par l'estimation des paramètres de la forme fonctionnelle. Par ailleurs, l'examen des fonctions d'autocorrélation et d'autocorrélation partielle des résidus après l'introduction du schéma autorégressif ($\rho = \rho_1, \rho_3, \rho_4, \rho_{12}$) en (b), de sorte que les délais les plus faibles sont affectés par les paramètres ρ_ℓ, montre que l'autocorrélation a été corrigée sauf pour le quatrième ordre. On se rappelle que la valeur du test statistique χ^2 pour la comparaison des modèles BC_* et BC_*AU s'élève à 29.27; ceci implique,avec les 4 degrés de liberté requis par les quatre éléments non-nuls du vecteur ρ, que l'hypothèse nulle $\rho = 0$ peut être rejetée avec un niveau de confiance extrêmement élevé. Aussi, il n'est pas surprenant que la structure soit modifiée. Il est notable, si nous considérons le modèle qui optimise L^{***} à savoir (d), que l'autocorrélation est corrigée pour le quatrième ordre. Ceci signifie que les tests statistiques-t ne sont pas biaisés vers le haut.

Nous devons ajouter qu'excepté pour le modèle $BC_{CC}AUHC_{1,2}$ (l'ordre 9 est alors égal à .29 (résultat non montré)), nous n'avons pas pu isoler un autre schéma autorégressif ce qui est très rassurant.

. Coefficient de détermination multiple.

TABLEAU 13

Fonctions d'autocorrélation et d'autocorrélation partielle, en différence 0 des résidus.

(a) $\lambda_y=\lambda_{x(1)}=\lambda_{x(2)}=\lambda_{x(3)}=\lambda_{x(4)}=\lambda_{x(5)}=1$; $\rho=0$; $\lambda_e=0$; $\delta=0$.

Délais					Autocorrélations							Ecart-type estimé pour la ligne
1-12	.16	-.02	-.19	-.19	-.05	.08	-.07	-.09	-.17	.00	.00	.38	(.08)
13-24	.07	-.00	-.12	-.08	-.02	.03	-.10	-.06	-.20	-.07	.03	.23	(.09)
25-36	.03	-.03	-.10	-.00	.02	.08	.01	-.05	-.08	-.06	.01	.18	(.10)
37-48	.02	.01	-.07	.07	.07	.00	-.05	-.17	-.08	-.02	.05	.19	(.11)
49-60	.01	-.01	-.06	-.03	.05	-.01	-.12	-.16	-.05	.02	.11	.20	(.11)
					Autocorrélations partielles							
1-12	.16	-.04	-.19	-.14	-.01	.05	-.16	-.11	-.15	.02	-.08	.33	(.08)

(b) $\lambda_y=\lambda_{x(1)}=\lambda_{x(2)}=\lambda_{x(3)}=\lambda_{x(4)}=\lambda_{x(5)}=1$; $\hat{\rho}=\hat{\rho}_1, \hat{\rho}_3, \hat{\rho}_4, \hat{\rho}_{12}$; $\lambda_e=0$; $\delta=0$.

					Autocorrélations							
1-12	.04	.01	.00	-.16	-.00	.03	-.08	.03	-.12	.09	-.09	-.02	(.08)
13-24	.03	.04	.01	.03	-.04	-.05	-.13	-.01	-.21	-.02	.02	.05	(.08)
25-36	-.02	-.06	-.00	.00	-.01	.18	.05	.03	-.02	-.07	.04	.03	(.09)
37-48	.02	.02	-.12	.03	.03	.00	.02	-.13	-.01	-.01	.01	.06	(.09)
49-60	-.02	.02	-.05	-.11	-.01	-.00	-.09	-.08	.02	.02	.08	.11	(.09)
					Autocorrélations partielles							
1-12	.04	.01	.00	-.16	.01	.04	-.09	.02	-.12	.12	-.13	-.00	(.08)

UNIVERSITÉ DE MONTRÉAL

(c) $\hat{\lambda}_y=\hat{\lambda}_{x(1)}=\hat{\lambda}_{x(2)}=\hat{\lambda}_{x(3)}=\hat{\lambda}_{x(4)}=\hat{\lambda}_{x(5)}=.51$; $\rho=0$; $\lambda_e=0$; $\delta=0$.

Délais					Autocorrélations							Ecart-type estimé pour la ligne
1-12	.15	-.05	-.20	-.17	-.05	.07	-.09	-.11	-.16	-.00	-.01	.35	(.08)
13-24	.08	.02	-.11	-.07	-.04	.03	-.08	-.04	-.16	-.07	-.00	.20	(.09)
25-36	.06	.01	-.09	-.01	-.00	.06	.02	-.02	-.06	-.07	-.01	.15	(.10)
37-48	.03	.04	-.08	.05	.06	-.00	-.03	-.14	-.08	-.03	.04	.15	(.10)
49-60	.03	.02	-.06	-.03	.07	.02	-.09	-.17	-.05	.01	.10	.18	(.11)
					Autocorrélations partielles								
1-12	.15	-.07	-.19	-.13	-.03	.03	-.17	-.13	-.16	-.02	-.13	.29	(.08)

(d) $\hat{\lambda}_y=.57$, $\hat{\lambda}_{x(1)}=\hat{\lambda}_{x(3)}=\hat{\lambda}_{x(4)}=\hat{\lambda}_{x(5)}=.96$, $\hat{\lambda}_{x(2)}=.68$; $\hat{\rho}=\hat{\rho}_1$, $\hat{\rho}_3$, $\hat{\rho}_4$, $\hat{\rho}_{12}$; $\hat{\lambda}_{e1}$, $\hat{\lambda}_{e2}$; $\hat{\delta}_1$, $\hat{\delta}_2$.

					Autocorrélations								
1-12	-.09	.05	-.04	-.07	.11	.06	-.14	.14	-.02	.07	-.03	.02	(.08)
13-24	-.07	.09	-.02	.07	-.07	-.02	-.09	-.03	-.16	.05	.03	.06	(.08)
25-36	-.05	-.03	-.02	-.01	-.05	.05	.02	.04	.05	-.08	.04	.07	(.09)
37-48	-.02	.10	-.12	-.01	.08	-.05	.05	.00	-.04	.02	.01	-.00	(.09)
49-60	.08	.07	-.10	.01	.01	.00	-.01	-.06	-.05	.09	-.02	.10	(.09)
					Autocorrélations partielles								
1-12	-.09	.04	-.04	-.08	.10	.08	-.15	.12	.04	.04	-.04	.06	(.08)

UNIVERSITÉ DE MONTRÉAL

Le tableau 14 présente les résultats obtenus pour les trois mesures du coefficient de corrélation multiple R^2. (Une description exhaustive des trois mesures est fournie dans l'appendice C).

Les valeurs générées pour le coefficient de détermination "intuitif" étant très élevées (supérieures à .90), elles prouvent la qualité de l'ajustement du modèle de demande à la réalité statistique.

La première mesure de R^2 est soit inférieure, soit supérieure aux deux autres mesures. Cependant, la première mesure étant plus convaincante de par sa définition de la mesure de l'erreur, elle doit être considérée.

Il est notable que les principaux gains au niveau du premier coefficient de détermination sont réalisés lorsque nous libérons les contraintes sur les éléments du vecteur ρ; l'introduction de la formulation générale de l'hétéroscédasticité n'apporte aucune modification appréciable à ce niveau.

Cependant, une analyse de $R^2(E(y))$ au niveau des modèles $BC_A AUHG_1$, $BC_A AUHG_2$ versus $BC_{AA} AUHG_{1,2}$ nous permet de constater que le coefficient de détermination du dernier modèle est plus faible (.924) que celui des deux premiers modèles (.931) et (.934), même si la variance résiduelle est plus faible pour le modèle $BC_{AA} AUHG_{1,2}$ (.012) lorsque comparée au modèle $BC_A AUHG_1$ (.022) et équivalente lorsque comparée au modèle $BC_A AUHG_2$.

4.3.2. Effets de la forme fonctionnelle de la partie systématique et de la spécification de l'erreur sur la signification des variables.

Le choix de la forme fonctionnelle peut s'avérer critique pour le niveau de signification des variables dans l'équation. En effet, une spécification erronée de la forme fonctionnelle de la partie systématique ou de la partie stochastique du modèle ou encore des deux parties prises conjointement peut mener à l'inclusion ou à l'exclusion de variables explicatives.

TABLEAU 14

Valeurs des coefficients paramétriques, des élasticités, des tests statistiques, de L*** pour divers modèles.

		I BC*		II BC*AU		III BC*		IV BC*AU		V BC*AUHG$_2$	
	Code des paramètres	c	t	c	t	c	t	c	t	c	t
	λ_y	0.00		0.00		1.00		1.00		1.00	
	$\lambda_{x(1)}$	0.00		0.00		1.00		1.00		1.00	
	$\lambda_{x(2)}$	0.00		0.00		1.00		1.00		1.00	
	$\lambda_{x(3)}$	0.00		0.00		1.00		1.00		1.00	
	$\lambda_{x(4)}$	0.00		0.00		1.00		1.00		1.00	
	$\lambda_{x(5)}$	0.00		0.00		1.00		1.00		1.00	
	ρ_1	0.00		.10	1.33	0.00		.17	2.26	.25	2.92
	ρ_3	0.00		-.34	-3.88	0.00		-.32	-3.64	-.19	-1.82
	ρ_4	0.00		-.10	-1.22	0.00		-.07	-.87	-.06	-.58
	ρ_{12}	0.00		.33	4.36	0.00		.41	5.19	.47	6.84
	λ_{e1}	0.00		0.00		0.00		0.00		0.00	
	λ_{e2}	0.00		0.00		0.00		0.00		1.08	1.89 .13
	δ_1	0.00		0.00		0.00		0.00		0.00	
	δ_2	0.00		0.00		0.00		0.00		.12	.98
Code	Nom des variables	n	t	n	t	n	t	n	t	n	t
FSCD	Tarif réel étudiant	-.23	-1.30	-.40	-2.03	-.38	-2.31	-.44	-2.46	-.41	-2.73
NC	Nombre d'automobiles	.11	.21	-.89	-1.05	-.92	-1.38	-1.16	-1.80	-1.04	-1.66
P	Ind. des prix à la cons.	-1.59	-1.08	1.55	.80	-1.08	-.71	.54	.28	-.16	-.12
RAWEM	Salaire réel moyen	-.20	-.28	-.55	-.77	.91	1.49	.65	1.26	.82	1.06
TW	Temps d'attente du t.e.c.	-.36	-.84	-.25	-.62	-.07	-.33	-.21	-.61	-.33	-.90
TT	Temps de trans. en t.e.c.	.74	1.64	.14	.30	.30	.80	-.16	-.37	-.06	-.09
T2T	Temps de trans. en auto	1.38	1.29	.36	.29	1.00	1.27	.43	.39	.82	.69
TLK	Inc. dû à la température	.02	2.28	.02	1.75	.07	3.17	.07	2.50	.02	.70
RFK	Inc. dû à la pluie	-.01	-1.26	-.01	-1.28	-.01	-.87	-.01	-1.03	-.01	-1.15
SFK	Inc. dû à la neige	-.01	-2.44	-.00	-.88	-.03	-3.25	-.01	-1.79	-.01	-1.17
CSFMR	Inc. dû à la neige acc.	.01	1.53	.00	.56	.02	1.79	.01	.87	.01	.95
RSCMP	Ind. de présence scolaire	.17	12.59	.15	7.06	.44	11.84	.34	5.67	.34	6.52
SHONF	Ind. de dépl.-magasinage	.12	2.41	.11	1.86	.14	3.74	.11	2.36	.09	1.58
EX7	Expo'67	.00	5.93	.00	5.47	.01	5.38	.01	4.91	.01	4.04
EX8	Expo'68	.00	2.00	.00	1.34	.00	1.42	.00	1.21	.00	.63
STR	Grève à la C.T.C.U.M.	-.00	-13.94	-.00	-14.84	-.01	-11.52	-.01	-12.60	-.01	-11.95
TEND	Trend	.21	1.15	.40	1.58	.54	2.16	.58	2.26	.57	1.77
ANT	Ventes anticipées	.00	2.19	.00	.78	.00	4.14	.00	1.37	.00	1.17
ANTL1	Ventes ant. retardées	-.00	-1.16	-.00	-.78	-.00	-1.24	-.00	- 56	-.00	-.38
ZSC	Double comptage zonal	.02	2.33	.02	2.78	.07	3.57	.05	3.56	.05	3.23
WD	Jours ouvrables/mois	.77	3.34	.57	1.96	.87	3.89	.67	2.12	.72	2.29
SAT	Samedis/mois	.11	1.30	.06	.60	.11	1.51	.05	.62	.07	84
SHD	Dim. & jrs. fériés/mois	.07	1.02	.06	.79	.06	.94	.05	.77	.10	1.30
CHEAT2	Contrôle-carte d'écolier	-.00	-2.20	-.00	-.56	-.01	-1.16	-.00	-.30	-.01	-.57
STRUCT1	Restructuration-lignes	.02	1.21	.00	.16	.01	.35	-.03	-.81	-.02	-.45
CARDEX	Expiration-carte étud.	.00	.15	-.00	-.75	.00	.40	-.00	-.51	-.00	-.43
C	Constante	-.28	-.19	.60	.36	-1.23	-2.12	-.72	-.82	-.71	-.89
$R^2(E(y))$		.902		.922		.907		.934		.932	
$R^2(E(y^{**}))$		.927		.891		.907		.850		.858	
$R^2(\bar{y}^{**})$		.927		.891		.907		.850		.858	
Variance résiduelle		.007		.005		.060		.043		.012	
L***		235.84		257.36		237.10		266.82		277.28	

TABLEAU 14 (suite)

Valeurs des coefficients paramétriques, des élasticités, des tests statistiques, de L^{***} pour divers modèles.

		VI BC_A		VII BC_AAU		VIII BC_AAUHG_1		IX BC_AHG_2		X BC_AAUHG_2	
	Code des paramètres	C	t	C	t	C	t	C	t	C	t
	λ_y	.51	10.51 -10.07	.72	13.04 -4.99	.75	3.55 -1.17	.48	4.20 -4.50	1.004	5.94 .03
	$\lambda_{x(1)}$	.51		.72		.75		.48		1.004	
	$\lambda_{x(2)}$	.51		.72		.75		.48		1.004	
	$\lambda_{x(3)}$	.51		.72		.75		.48		1.004	
	$\lambda_{x(4)}$	.51		.72		.75		.48		1.004	
	$\lambda_{x(5)}$	.51		.72		.75		.48		1.004	
	ρ_1	0.00		.14	1.82	.11	1.45	0.00		.25	2.94
	ρ_3	0.00		-.34	-3.73	-.33	-3.36	0.00		-.19	-1.77
	ρ_4	0.00		-.09	-1.08	-.10	-1.20	0.00		-.06	-.57
	ρ_{12}	0.00		.38	4.66	.39	4.67	0.00		.47	6.87
	λ_{e1}	0.00		0.00		-.46	-.03 -.10	0.00		0.00	
	λ_{e2}	0.00		0.00		0.00		2.14	1.83 .98	1.07	1.30 .09
	δ_1	0.00		0.00		.17	.11	0.00		0.00	
	δ_2	0.00		0.00		0.00		.01	.38	.12	.60
Code	Nom des variables	n	t	n	t	n	t	n	t	n	t
FSCD	Tarif réel étudiant	-.29	-1.10	-.42	-2.33	-.42	-1.77	-.39	-2.06	-.41	-2.09
NC	Nombre d'automobiles	-.71	-.81	-1.14	-1.66	-1.10	-1.58	-.91	-1.29	-1.04	-1.65
P	Ind. des prix à la cons.	-1.60	-.74	.58	.30	.62	.32	.17	.12	-.17	-.12
RAWEM	Salaire réel moyen	.21	.28	.27	.49	.25	.45	.58	.68	.82	1.05
TW	Temps d'attente du t.e.c.	-.39	-.97	-.28	-.79	-.24	-.65	-.19	-.53	-.33	-.89
TT	Temps de trans. en t.e.c.	.51	.69	-.10	-.25	-.12	-.27	.21	.31	-.06	-.09
TZT	Temps de trans. en auto	1.94	1.34	.71	.60	.64	.52	.74	.61	.82	.63
TLK	Inc. dû à la température	.03	1.83	.05	2.58	.05	2.07	.03	1.72	.02	.67
RFK	Inc. dû à la pluie	-.01	-.80	-.01	-1.05	-.01	-.98	-.01	-.68	-.01	-1.14
SFK	Inc. dû à la neige	-.02	-2.27	-.01	-1.50	-.01	-1.42	-.01	-1.12	-.01	-1.16
CSFMR	Inc. dû à la neige acc.	.01	1.20	.01	.60	.01	.57	.00	.67	.01	.93
RSCMP	Ind. de présence scolaire	.27	7.82	.27	6.40	.28	4.16	.27	7.93	.34	4.76
SHONF	Ind. de dépl.-magasinage	.13	2.66	.11	2.37	.12	2.19	.14	2.59	.09	1.57
EX7	Expo'67	.01	6.25	.01	5.45	.01	3.61	.01	5.10	.01	3.78
EX8	Expo'68	.00	1.45	.00	1.33	.00	1.32	.00	1.26	.00	.63
STR	Grève à la C.T.C.U.M.	-.01	-4.78	-.01	-11.19	-.01	-4.74	-.01	-4.69	-.01	-6.01
TEND	Trend	.60	1.78	.60	2.22	.58	1.65	.39	1.15	.57	1.77
ANT	Ventes anticipées	.00	.00	.00	1.22	.00	1.07	.00	.00	.00	1.05
ANTL1	Ventes ant. retardées	-.00	-.00	-.00	-.67	-.00	-.62	-.00	-.00	-.00	-.38
ZSC	Double comptage zonal	.03	1.88	.04	3.15	.04	2.39	.03	2.14	.05	2.91
WD	Jours ouvrables/mois	.81	2.85	.66	2.11	.69	2.08	.89	3.09	.71	2.24
SAT	Samedis/mois	.13	1.46	.05	.64	.06	.67	.13	1.47	.07	.83
SHD	Dim. & jrs. fériés/mois	.07	1.02	.06	.77	.06	.79	.08	1.19	.10	1.30
CHEAT2	Contrôle-carte d'écolier	-.00	-1.30	-.00	-.27	-.00	-.26	-.01	-1.24	-.01	-.57
STRUCT1	Restructuration-lignes	.02	.44	-.02	-.68	-.02	-.70	.00	.06	-.02	-.45
CARDEX	Expiration-carte étud.	.00	.19	-.00	-.53	-.00	-.56	-.00	-.34	-.00	-.42
C	Constante	-.56	-.46	-.41	-.42	-.45	-.42	-1.27	-1.13	-.71	-.89
$R^2(E(y))$		.908		.933		.934		.904		.931	
$R^2(E(y^{**}))$		.919		.869		.870		.961		.857	
$R^2(\bar{y}^{**})$		.919		.869		.870		.961		.857	
Variance résiduelle		.019		.022		.022		.006		.012	
L^{***}		241.19		268.00		268.07		253.72		277.28	

TABLEAU 14 (suite)

Valeurs des coefficients paramétriques, des élasticités, des tests statistiques, de L*** pour divers modèles.

		XI BC_AAUHC_2		XII BC_BHG_2		XIII BC_BAUHG_2		XIV BC_CAUHG_1		XV BC_CHG_2	
Code des paramètres		C	t	C	t	C	t	C	t	C	t
	λ_y	1.14	17.94 2.21	.44	3.83 -4.82	1.08	7.35 .54	.60	3.99 -2.70	.42	3.46 -4.76
	$\lambda_{x(1)}$	1.14		.63	1.08 -.62	2.46	1.68 1.00	4.61	-2.70 2.11	2.99	1.31 .87
	$\lambda_{x(2)}$	1.14		.63		2.46		4.61		2.99	
	$\lambda_{x(3)}$	1.14		.63		2.46		4.61		2.99	
	$\lambda_{x(4)}$	1.14		.63		2.46		1.23	3.83 .71	.74	1.73 -.61
	$\lambda_{x(5)}$	1.14		.63		2.46		4.61		2.99	
	ρ_1	-.09	-.78	0.00		.34	4.75	.09	1.06	0.00	
	ρ_3	-.39	-5.69	0.00		-.24	-2.73	-.51	-6.37	0.00	
	ρ_4	-.27	-2.53	0.00		.04	.52	-.09	-1.05	0.00	
	ρ_{12}	.15	2.65	0.00		.53	9.45	.28	3.97	0.00	
	λ_{e1}	0.00		0.00		0.00		2.78	.40 .26	0.00	
	λ_{e2}	0.00		2.17	1.78 .96	.91	1.53 -.16	0.00		2.02	1.69 .65
	δ_1	0.00		0.00		0.00		-.04	-.14	0.00	
	δ_2	2.00		.01	.37	.17	-.89	0.00		.02	.38
Code	Nom des variables	n	t	n	t	n	t	n	t	n	t
FSCD	Tarif réel étudiant	-.37	-4.64	-.39	-1.76	-.44	-1.17	-.35	-1.28	-.35	-.92
NC	Nombre d'automobiles	-1.35	-2.86	-1.04	-1.25	-1.05	-.53	-1.42	-.48	-1.32	-.37
P	Ind. des prix à la cons.	2.09	2.59	.20	.14	-.01	-.00	.96	.50	1.00	.45
RAWEM	Salaire réel moyen	.31	.44	.59	.62	1.28	.38	.25	.21	.52	.18
TW	Temps d'attente du t.e.c.	.16	.99	-.10	-.19	-.48	-.49	-.47	-.39	-.38	-.28
TT	Temps de trans. en t.e.c.	-.78	-1.48	.20	.31	-.16	-.35	-.18	-.52	.15	.27
T2T	Temps de trans. en auto	-.22	-.68	.78	.64	.68	.67	3.95	1.47	2.34	.94
TLK	Inc. dû à la température	.08	3.41	.03	1.68	-.01	-.19	.05	2.94	.02	1.38
RFK	Inc. dû à la pluie	.00	.13	-.01	-.74	-.01	-1.24	-.01	-1.26	-.01	-.73
SFK	Inc. dû à la neige	-.01	-1.61	-.01	-1.21	-.01	-1.09	-.00	-.87	-.01	-1.10
CSFMR	Inc. dû à la neige acc.	.01	.62	.00	.67	.01	.97	-.00	-.13	.00	.50
RSCMP	Ind. de présence scolaire	.46	7.96	.26	8.05	.30	4.34	.23	6.46	.25	8.74
SHONF	Ind. de dépl.-magasinage	.18	3.46	.14	1.73	.07	.66	.09	2.15	.12	1.81
EX7	Expo'67	.01	3.91	.01	4.64	.01	3.51	.00	3.74	.00	4.10
EX8	Expo'68	.00	.67	.00	1.29	.00	.90	.00	1.98	.00	1.29
STR	Grève à la C.T.C.U.M.	-.01	-4.10	-.01	-4.63	-.01	-6.65	-.01	-6.07	-.01	-4.20
TEND	Trend	.45	2.52	.44	.72	.49	.28	.71	1.14	.47	.92
ANT	Ventes anticipées	.00	.95	.00	.00	.00	1.54	.00	1.43	.00	.00
ANTL1	Ventes ant. retardées	-.00	-.78	-.00	-.00	-.00	-.50	-.00	-1.26	-.00	-.00
ZSC	Double comptage zonal	.06	1.99	.03	2.01	.06	2.96	.03	2.74	.03	2.10
WD	Jours ouvrables/mois	.79	3.70	.89	2.04	.47	.81	.50	.74	.89	.59
SAT	Samedis/mois	.14	2.23	.12	.80	.03	.28	.04	.32	.13	.28
SHD	Dim. & jrs. fériés/mois	.07	1.27	.08	.77	.04	.35	.01	.11	.07	.25
CHEAT2	Contrôle-carte d'écolier	.00	.10	-.00	-.93	-.00	-.14	-.00	-.19	-.00	-1.03
STRUCT1	Restructuration-lignes	-.10	-1.68	.00	.05	-.05	-1.19	-.04	-2.12	-.00	-.19
CARDEX	Expiration-carte étud.	-.01	-.99	-.00	-.30	-.00	-.59	-.00	-.25	-.00	-.38
C	Constante	-.73	-1.33	-1.18	-1.20	-.32	-.52	-.61	-1.23	-.79	-1.44
$R^2(E(y))$		.916		.904		.932		.936		.905	
$R^2(E(y^{**}))$		.990		.962		.816		.926		.963	
$R^2(\hat{y}^{**})$		.990		.962		.816		.926		.963	
Variance résiduelle		.001		.006		.013		.019		.006	
L***		242.17		253.86		277.69		273.05		255.04	

TABLEAU 14 (suite)

Valeurs des coefficients paramétriques, des élasticités, des tests statistiques, de L*** pour divers modèles.

		XVI $BC_C AUHG_2$		XVII $BC_E HG_2$		XVIII $BC_{AA} AUHG_{1,2}$		XIX $BC_{AA} AUHC_{1,2}$		XX $BC_{CC} AUHG_{1,2}$	
	Code des paramètres	C	t	C	t	C	t	C	t	C	t
	λ_y	.80	5.57 -1.41	.32	2.88 -4.76	.92	6.84 -.61	1.22	20.55 3.66	.57	4.21 -3.24
	$\lambda_{x(1)}$	3.27	1.84 1.28	9.05	4.17 3.70	.92		1.22		.96	1.51 -.07
	$\lambda_{x(2)}$	3.27		7.07	.82 .70	.92		1.22		.96	
	$\lambda_{x(3)}$	3.27		5.20	1.53 1.23	.92		1.22		.96	
	$\lambda_{x(4)}$	1.17	2.89 .41	.09	.33 -3.25	.92		1.22		.68	2.07 -.97
	$\lambda_{x(5)}$	3.27		-.87	-.19 -.40	.92		1.22		.96	
	ρ_1	.06	.74	0.00		.53	9.24	.35	3.18	.60	7.94
	ρ_3	-.40	-4.82	0.00		-.59	-8.17	-.30	-3.86	-.41	-4.39
	ρ_4	-.13	-1.49	0.00		.35	5.09	-.02	-.21	.11	1.21
	ρ_{12}	.31	4.02	0.00		.32	4.87	-.09	-1.81	.32	4.26
	λ_{e1}	0.00		0.00		5.38	9.10 7.41	0.00		4.82	7.71 6.11
	λ_{e2}	1.38	1.19 .33	1.70	1.23 .50	.36	.88 -1.58	0.00		1.89	1.72 .81
	δ_1	0.00		0.00		-.03	-1.43	2.00		-.05	-1.39
	δ_2	.05	.42	.03	.34	.49	1.56	2.00		.02	.42
Code	Nom des variables	n	t	n	t	n	t	n	t	n	t
FSCD	Tarif réel étudiant	-.43	-1.15	-.15	-.80	-.58	-2.94	-.20	-1.23	-.53	-2.08
NC	Nombre d'automobiles	-1.53	-.49	-.35	-.32	-2.84	-7.57	-1.05	-1.36	-2.76	-1.37
P	Ind. des prix à la cons.	1.72	.79	-.48	-.27	1.52	1.25	1.77	1.47	1.18	.70
RAWEM	Salaire réel moyen	.49	.26	.71	.05	2.13	6.44	-.41	-.58	1.72	.67
TW	Temps d'attente du t.e.c.	-.32	-.41	-.88	-.19	-.60	-1.99	.07	.30	-.61	-.93
TT	Temps de trans. en t.e.c.	-.21	-.44	.19	.28	-.17	-.88	-1.63	-2.93	-.12	-.51
T2T	Temps de trans. en auto	2.11	.87	7.36	2.38	.94	1.56	.38	.80	1.57	1.71
TLK	Inc. dû à la température	.05	2.63	.02	1.59	.03	1.30	.05	2.26	.00	.30
RFK	Inc. dû à la pluie	-.01	-.73	-.00	-.75	-.01	-1.14	-.05	-5.13	-.01	-1.98
SFK	Inc. dû à la neige	-.00	-.54	-.01	-1.11	-.00	-.51	.00	.50	-.00	-.49
CSFMR	Inc. dû à la neige acc.	.00	.05	.00	.68	-.00	-.03	.01	.75	-.00	-.09
RSCMP	Ind. de présence scolaire	.30	6.44	.21	6.81	.26	5.54	.47	8.21	.20	6.04
SHONF	Ind. de dépl.-magasinage	.12	2.02	.14	1.91	.13	3.23	.10	1.96	.16	2.14
EX7	Expo'67	.00	3.70	.00	3.67	.00	2.03	.01	1.74	.00	1.92
EX8	Expo'68	.00	1.39	.00	.87	.00	.87	.00	.24	.00	.83
STR	Grève à la C.T.C.U.M.	-.01	-6.19	-.01	-5.41	-.01	-6.93	-.01	-3.93	-.01	-7.70
TEND	Trend	.58	.97	.41	2.00	1.03	3.90	.60	2.05	1.06	1.28
ANT	Ventes anticipées	.00	1.35	.00	.00	.00	2.35	.00	1.30	.00	1.32
ANTL1	Ventes ant. retardées	-.00	-.60	-.00	-.00	-.00	-.98	-.00	-.44	-.00	-.90
ZSC	Double comptage zonal	.04	3.06	.04	2.46	.06	3.62	.05	1.47	.04	3.78
WD	Jours ouvrables/mois	.66	.72	.80	.30	.43	2.21	.70	5.17	.52	1.60
SAT	Samedis/mois	.06	.34	.15	.14	.01	.16	.18	4.14	.01	.23
SHD	Dim. & jrs. fériés/mois	.04	.29	.08	.13	.03	.61	-.00	-.02	.07	.73
CHEAT2	Contrôle-carte d'écolier	-.00	-.24	-.01	-1.31	-.01	-1.40	.00	.04	-.00	-.92
STRUCT1	Restructuration-lignes	-.04	-1.38	-.02	-.97	-.02	-1.40	-.23	-2.95	-.01	-.85
CARDEX	Expiration-carte étud.	-.00	-.66	.00	.05	.00	.37	-.01	-1.07	.00	.08
C	Constante	-.77	-1.55	-3.21	-.59	-1.11	-1.89	.50	.81	-.92	-1.23
$R^2(E(y))$		.935		.915		.924		.876		.924	
$R^2(E(y^{**}))$		.943		.961		.999		.994		.999	
$R^2(\hat{y}^{**})$		.943		.961		.999		.994		.999	
Variance résiduelle		.011		.005		.012		.000		.011	
L***		278.34		260.00		295.77		216.97		296.81	

Il importe de préciser que les différentes applications inhérentes à la fonction de demande ne nous permettent pas d'effectuer une comparaison rigoureuse au niveau de la signification des variables explicatives puisqu'une modification de la structure du terme de l'erreur affecte la variance de l'erreur et par conséquent, les statistiques-t de student. Néanmoins, nous croyons opportun d'établir des comparaisons (à ce niveau) entre différentes paires d'hypothèses, puisqu'elles mettront en relief les difficultés reliées à une mauvaise spécification de la forme fonctionnelle et/ou du terme de l'erreur.

Pour justifier ces assertions, nous examinerons à partir des résultats présentés dans le tableau 14, dans quelle mesure la signification des variables peut être affectée si

1) nous imposons une spécification arbitraire sur la partie fixe,
2) nous modifions les hypothèses de la variable aléatoire tout en conservant une spécification contraignante de la partie systématique,
3) nous modifions les hypothèses du terme de l'erreur tout en laissant les paramètres de la partie fixe s'estimer librement,
4) nous relâchons les contraintes de la partie systématique tout en maintenant inchangée la structure de la matrice variance-covariance des erreurs.

Les modèles BC_*AU ($\lambda_y = \lambda_{x(1)} = \lambda_{x(2)} = \lambda_{x(3)} = \lambda_{x(4)} = \lambda_{x(5)} = 0; \hat{\rho}$), BC_*AU ($\lambda_y = \lambda_{x(1)} = \lambda_{x(2)} = \lambda_{x(3)} = \lambda_{x(4)} = \lambda_{x(5)} = 1; \hat{\rho}$) et BC_AAU ($\hat{\lambda}_y = \hat{\lambda}_{x(1)} = \hat{\lambda}_{x(2)} = \hat{\lambda}_{x(3)} = \hat{\lambda}_{x(4)} = \hat{\lambda}_{x(5)} = .72; \hat{\rho}$) constituent les éléments explicatifs de notre première approche. Dans le premier modèle, 11 des 27 variables explicatives (y compris la constante) sont significativement différentes de zéro à un niveau égal ou supérieur à 0.10^4, comparativement à 12 pour les deux autres modèles.

4.- A un niveau de signification de 10%, la valeur critique du student pour ∞ degrés de liberté s'élève à 1.282.

Etant donné le nombre élevé de variables explicatives présentes dans le modèle, nous limiterons la comparaison des valeurs statistiques aux variables de prix, de revenu et de niveau de service.

Nos résultats montrent que si la forme logarithmique était imposée, seule la variable tarif FSCD serait significative alors que si la forme linéaire était imposée ou si une forme de puissance 3/4 était estimée, la variable nombre d'automobiles NC serait également significative.

Notre seconde approche est illustrée sur la base des modèles BC_*, BC_*AU et BC_*AUHG_2, lesquels sont assujettis à la forme linéaire. Les colonnes III, IV et V du tableau 14 indiquent que 17, 12 et 10 variables sont significatives pour chacun des trois modèles. Par ailleurs, la valeur numérique des tests statistiques est généralement plus élevée pour le modèle BC_* que pour le modèle BC_*AU. La différence dans les valeurs numériques entre les deux modèles peut s'expliquer en grande partie par le fait que dans le linéaire, l'absence de correction pour l'autocorrélation des erreurs résiduelles tend à biaiser les écarts-types vers le bas et par conséquent, les tests-t de student vers le haut.

Par analogie avec un schéma autorégressif qui incorpore un ordre simple positif, la correction pour l'autocorrélation des erreurs résiduelles à partir d'une structure d'autocorrélation d'ordre multiple fait diminuer la valeur numérique des tests statistiques. (Dans le non-linéaire, la comparaison des modèles BC_A versus BC_AAU, BC_BHG_2 versus BC_BAUHG_2 nous permet de dégager une conclusion similaire).

Nous pouvons noter,par ailleurs, que la sophistication de la spécification aléatoire (comparaison des modèles BC_*AU et BC_*AUHG_2) tend à faire augmenter le niveau de signification des variables tarif FSCD, temps d'attente TW

et temps de déplacement du mode concurrent T2T. Les deux dernières variables ne sont pas cependant significatives au niveau considéré (90%).

Notre troisième approche, quant à elle, consiste à examiner l'impact de la forme fonctionnelle sur la signification des variables à partir des modèles BC_AAU, BC_AAUHG_1, BC_AAUHG_2, BC_AAUHC_2 et $BC_{AA}AUHG_{1,2}$ pour lesquels 12, 12, 10, 16 et 17 variables sont respectivement significatives. Les résultats dégagés par ces modèles sont extrêmement révélateurs de l'utilité de bien spécifier le modèle. En effet, ils montrent très clairement que la signification statistique des variables est en général plus élevée pour le modèle qui optimise la fonction de vraisemblance, à savoir le modèle[5] $BC_{AA}AUHG_{1,2}$. Plus important encore, nos résultats montrent que l'utilisation irrationnelle des autres modèles entraînerait l'élimination des variables revenu RAWEM, temps d'attente TW et temps de déplacement en auto T2T dans l'équation de demande.

Il est par ailleurs assez surprenant de constater que la correction par la formulation classique de l'hétéroscédasticité produit un meilleur effet sur la signification des variables que la correction par la formulation générale de l'hétéroscédasticité, lorsque nous considérons les variables *SHONF* et *TEND* isolément dans $f(X_t)$. En effet, si nous considérons les modèles BC_AAUHG_1 (BC_AAUHG_2) et BC_AAUHC_1 (BC_AAUHC_2), nous trouvons que 19 (18) des 26 variables explicatives possèdent un niveau de signification plus élevé dans le modèle qui incorpore la formulation classique de l'hétéroscédasticité. (Les résultats pour le modèle BC_AAUHC_1 ne sont pas présentés.) Cependant, lorsque les variables *SHONF* et *TEND* sont introduites comme sources de l'hétéroscédasticité, 16 des 26 variables explicatives possèdent des valeurs statistiques plus élevées

---

5.- Même si la valeur obtenue pour L^{***} confère l'optimalité du modèle $BC_{CC}AUHG_{1,2}$ (comme nous l'avons noté antérieurement), nous pouvons constater que ce modèle perd beaucoup au niveau de la signification des variables explicatives. Une explication pertinente à ce problème est soulevée dans la quatrième approche.

avec la formulation générale de l'hétéroscédasticité. Il ne semble donc pas y avoir de consensus établi ex-ante sur le fait que la valeur numérique du test statistique-t soit plus grande lorsque nous utilisons la formulation générale de l'hétéroscédasticité que lorsque nous utilisons la formulation classique de l'hétéroscédasticité.

Dans la dernière approche, nous allons considérer l'impact sur la signification des variables lorsqu'il y a relaxation des contraintes sur la partie fixe et une hypothèse spécifique sur le terme de l'erreur. Il ressort à travers les résultats présentés que le nombre de variables significatives est, règle générale, de moins en moins imposant au fur et à mesure que nous relâchons les contraintes sur la partie fixe du modèle. Pour mettre en relief cette affirmation, nous allons considérer le modèle BC_SHG_2 pour les séries * (λ = 1), A, B, C et E. Pour chacune des séries, le nombre de variables statistiquement significatives à un niveau d'au moins 90% s'établit respectivement à 13 (résultats non montrés), 10, 9, 8 et 9. (Nous pouvons également constater à travers les diverses applications de ce modèle que le niveau de signification des variables semble décroître, règle générale, lorsque nous généralisons la paramétrisation Box-Cox. Considérons les variables tarif FSCD et revenu RAWEM. Pour la variable FSCD, la signification statistique passe de -2.06 dans la série A à -1.76 dans B, à -.92 dans C et à -.80 dans E, tandis que pour la variable RAWEM, elle passe de .68 dans A, à .62 dans B, à .18 dans C et à .05 dans E.) L'explication la plus logique attribuée à cet état de fait est, selon le professeur Dagenais, reliée à la variante qui existe entre la variance conditionnelle et la variance marginale de β.

D'après la loi conditionnelle, nous pouvons écrire:

$$\text{Var conditionnelle } (\hat{\beta} \mid \hat{\lambda}_{xk} = \hat{\lambda}_y) = E[\hat{\beta} - E(\hat{\beta} \mid \hat{\lambda}_{xk} = \hat{\lambda}_y)]^2.$$

Or, si nous relâchons la contrainte d'égalité entre $\hat{\lambda}_{xk}$ et $\hat{\lambda}_y$, la loi

marginale nous permet d'écrire:

$$\text{Var marginale } (\hat{\beta}) = E[\hat{\beta} - E(\hat{\beta})]^2.$$

Or, si la variance marginale est plus grande que la variance conditionnelle, cela biaise la valeur du test de student vers le bas. Nos résultats tendent à confirmer cette assertion.

Il est à noter que les valeurs obtenues pour les tests-t de student de certaines variables explicatives et des coefficients paramétriques ne sont pas tout à fait exactes (elles sont surestimées, surtout celles associées au paramètre λ_y) à cause du mauvais calcul de la matrice de covariance, tel qu'indiqué dans le chapitre II.

Après un examen approfondi des effets de la forme fonctionnelle des parties fixe et aléatoire du modèle sur la signification des variables, nous allons, dans la prochaine étape, dénoter l'influence de ces éléments sur les signes des paramètres.

Dans l'ensemble, nous pouvons affirmer que les signes des coefficients de régression répondent assez bien aux anticipations du modèle théorique, quelque soit la détermination de la forme fonctionnelle.

Outre les variables de prix, de revenu et de niveau de service, les variables de température et de décision administrative exhibent également à quelques reprises des signes incompatibles avec la théorie économique. Les signes attendus pour le dernier groupe de variables (identifié par les variables CHEAT2, STRUCT1, CARDEX) sont négatifs (puisque les contrôles administratifs et la réorganisation des lignes d'autobus suite à l'implantation du métro, furent effectués au détriment des écoliers) alors que pour les variables de confort, il y a une forte présomption à croire que les températures excessivement chau-

des ou froides TLK et l'accumulation de neige au sol CSFMR vont augmenter le nombre d'usagers du transport en commun et que la pluie et les chutes de neige vont détourner les clients potentiels de ce mode de transport.

Procédons maintenant à une étude plus exhaustive des variables de prix, de revenu et de niveau de service.

Les résultats révèlent, pour les divers modèles estimés, que dans le groupe des variables de prix, seul le tarif présente toujours le signe anticipé (négatif). En effet, lorsque la forme logarithmique associée à des hypothèses contraignantes du terme de l'erreur est imposée, les variables nombre d'automobiles NC et indice des prix à la consommation des biens autres que le transport P sont affectées du mauvais signe. De même, la variable P est affectée d'un signe erroné lorsque la forme puissance "racine carrée" est déterminée librement (mais accompagnée d'hypothèses restrictives sur la partie stochastique) ou encore lorsque la forme linéaire imposée ou déterminée librement (séries A et B) est liée à des hypothèses contraignantes ($\rho = 0$; $\lambda_e = 0$; $\delta = 0$) ou à des hypothèses moins restrictives ($\hat{\rho}_1$, $\hat{\rho}_3$, $\hat{\rho}_4$, $\hat{\rho}_{12}$; $\lambda_{e1} = 0$, $\hat{\lambda}_{e2}$; $\delta_1 = 0$, $\hat{\delta}_2$) sur la variable aléatoire. La variable revenu RAWEM, qui permet d'identifier la nature d'un mode spécifique, témoigne également de la nécessité de bien préciser la forme fonctionnelle. En effet, d'après les résultats induits, le transport public est considéré comme un bien inférieur si la forme logarithmique est assignée et comme un bien supérieur si les autres formes sont générées. Lorsque les variables RAWEM et P sont affectées simultanément d'un signe positif, cela implique que l'effet de substitution est plus grand que l'effet de revenu et signifie que le transport public est un substitut net pour les autres biens.

En ce qui a trait aux signes affectant les coefficients relatifs aux variables de temps, ils ne concordent également pas toujours avec nos attentes.

Parmi les trois variables de niveau de service (TW, TT, T2T), la variable temps de déplacement TT est indubitablement la plus tributaire du comportement de la forme fonctionnelle et de la spécification de l'erreur. Effectivement, nous remarquons qu'il y a un changement de signe pour cette variable toutes les fois où nous ne corrigeons pas pour l'autocorrélation des résidus; nous obtenons à ces occasions (sauf pour la forme logarithmique) un signe positif qui va dans la direction opposée à nos anticipations. Les variables temps d'attente TW et temps de déplacement en auto T2T présentent, quant à elles, des signes opposés et discordants lorsque nous introduisons la formulation classique de l'hétéroscédasticité.

Les coefficients autorégressifs d'ordre 1 et 12 exhibent les signes conformes à nos attentes (positifs), excepté pour le modèle BC_AAUHC_2 tandis que les paramètres d'ordre 3 et 4, pour lesquels les signes sont inconnus à priori, exhibent toujours des signes négatifs sauf pour le modèle BC_BAUHG_2 et les fois où nous impliquons *SHONF* et *TEND* conjointement dans $f(X_t)$; ρ_4 est alors positif.

Les brefs commentaires que nous pouvons émettre sur le signe des variables se résument au fait que malgré les faibles contrariétés qui se dégagent au niveau des signes des estimateurs, ceux-ci sont tout de même révélateurs de l'utilité de bien déterminer la forme fonctionnelle de la partie systématique mais surtout de la partie aléatoire du modèle.

Nous croyons que certains des résultats non espérés (mauvais signes et variables non significatives) peuvent être expliqués par la forte multicolinéarité qui existe entre certaines paires de variables explicatives. Soit le modèle BC_*AU ($\lambda = 0$). Dans ce modèle, les signes affectant les variables revenu RAWEM (négatif), temps de trajet en autobus TT (positif) et la variable dichotomique inhérente à la restructuration des lignes STRUCT1 (positif) peu-

vent sans doute être expliqués par l'intercorrélation élevée entre elles et certaines autres variables. Ainsi, la variable RAWEM est apparentée très fortement et positivement aux variables NC (.9765), P (.9756), TW (.9246), TEND (.8790), STRUCT1 (.8154) et négativement aux variables TT (-.7091), ZSC (-.7377). La variable TT est reliée négativement aux variables NC (-.7175), P (-.7753), RAWEM (-.7091), TW (-.6213), T2T (-.7378), STRUCT1 (-.9503). En ce qui a trait aux coefficients de corrélation entre la variable STRUCT1 et les variables autres que RAWEM et TT, ils s'élèvent pour les variables NC, P, TW et T2T respectivement à .8218, .8804, .7059 et à .7603.

Nous allons maintenant présenter dans le tableau 15, le nombre de paires distinctes de corrélation partielle existantes pour certains modèles.

TABLEAU 15

Nombre de paires de corrélation partielle évaluées à différents niveaux pour certains modèles.

Modèles	Niveaux					
	.90	.80	.70	.60	.50	Total
BC_*AU (λ=0)	8	7	12	16	12	55
BC_* (λ=1)	11	7	16	8	12	54
BC_*AU (λ=1)	9	7	11	11	13	51
BC_*AUHG_2 (λ=1)	0	7	14	10	24	55
BC_AAUHG_1	8	7	13	12	10	50
BC_B	11	6	16	10	11	54
BC_BAU	9	6	14	9	14	52
BC_BHG_2	10	21	11	4	12	58
BC_BHC_2	43	8	16	22	28	117
BC_BAUHG_2	1	4	3	8	9	25
BC_BAUHC_2	96	10	15	12	3	136
BC_CAUHG_1	8	11	12	5	16	52
BC_CHG_2	3	12	12	8	15	50
BC_CAUHG_2	6	12	18	11	12	59

Le tableau 15 révèle pour le modèle BC_*AU ($\lambda = 0$) par exemple, qu'il y a parmi les 325 paires de régresseurs, 55 paires de coefficients de corrélation partielle r_{ij}[6] plus grandes que .50; parmi celles-ci, 8 possèdent des coefficients plus grands que .90.

Les résultats présentés dans ce tableau semblent suggérer

1°) que la correction pour l'autocorrélation des erreurs résiduelles tend à faire diminuer légèrement la collinéarité entre les variables explicatives aussi bien dans le linéaire que dans le non-linéaire,

2°) que la transformation Box-Cox sur le terme de l'erreur introduit pour certains modèles de la collinéarité alors que pour certains autres elle tend à la faire diminuer,

3°) que la correction par la formulation classique de l'hétéroscédasticité engendre beaucoup plus de corrélation entre les variables explicatives que la correction par la formulation générale. Une comparaison des modèles BC_BHG_2 versus BC_BHC_2 et BC_BAUHG_2 versus BC_BAUHC_2 témoigne de cette constatation. Pour ces deux derniers modèles par exemple, nous calculons 96 paires de r_{ij} supérieures à .90 pour BC_BAUHC_2 contre une seule pour BC_BAUHG_2.

4.3.3. Effets de la forme fonctionnelle de la partie systématique et de la spécification de l'erreur sur les élasticités.

En dernière analyse, nous établirons des comparaisons entre les modèles à partir des élasticités dérivées de la demande.

Dans un contexte de modèle de demande, l'élasticité de la demande de

---

6.- Le coefficient de corrélation partielle r_{ij} est défini comme étant égal à:

$$r_{ij} = \frac{\Sigma\,(X_i^{**} - \overline{X}_i)\,(X_j^{**} - \overline{X}_j)}{\sqrt{\Sigma\,(X_i^{**} - \overline{X}_i)^2\;\Sigma\,(X_j^{**} - \overline{X}_j)^2}} \quad .$$

y par rapport à chacune des variables X_k est définie de la manière suivante:

$$\eta_k = \frac{\partial y}{\partial X_k} \cdot \frac{X_k}{y} \quad . \tag{53}$$

Dans le cas de la transformation Box-Cox, c'est-à-dire si $\lambda_y \neq 0$ et si $\lambda_{xk} \neq 0$, l'élasticité se transforme en:

$$\eta_k = \frac{\partial y}{\partial y^{(\lambda_y)}} \cdot \frac{\partial y^{(\lambda_y)}}{\partial X_k^{(\lambda_{xk})}} \frac{\partial X_k^{(\lambda_{xk})}}{\partial X_k} \cdot \frac{X_k}{y}$$

$$= \frac{1}{y^{\lambda_y - 1}} \beta_k X_k^{\lambda_{xk}-1} \frac{X_k}{y}$$

$$= \beta_k \frac{X_k^{\lambda_{xk}}}{y^{\lambda_y}} \quad , \tag{54}$$

laquelle est un cas spécial des élasticités dérivées dans le texte de Gaudry et Wills, (1978, pp.43-44).

L'élasticité de la demande est évaluée à la valeur moyenne des observations de y et de X_k (c'est-à-dire à $\bar{y}$ et à $\bar{X}_k$). Ainsi,

$$\bar{\eta}_k = \beta_k \frac{\bar{X}_k^{\lambda_{xk}}}{\bar{y}^{\lambda_y}} \quad . \tag{55}$$

Notons que si $\lambda_y = \lambda_{xk} = 1$, $\bar{\eta}_k$ devient égal à $\beta_k \frac{\bar{X}_k}{\bar{y}}$ (56)

et si $\lambda_y = \lambda_{xk} = 0$, $\bar{\eta}_k$ devient égal à β_k . (57)

Dans le cas où la forme logarithmique est imposée, nous avons le concept "d'élasticité constante" car les élasticités dérivées de la demande sont indépendantes du niveau des variables. En ce qui concerne les autres applications, les taux d'élasticité sont affectés par les valeurs des variables y et X_k et par les β_k de sorte que la mesure d'élasticité varie pour chacune des observations mensuelles.

Pour les buts expositoires, nous retiendrons uniquement les caractéristiques de prix, de revenu et de temps.

Lorsque nous comparons les élasticités de ces variables selon la forme fonctionnelle, nous pouvons noter que l'ordre de priorité et que la grandeur des élasticités des facteurs influençant les flux de passagers sont sensiblement différents.

Ainsi, sur la base des variables significatives, nous pouvons simplement souligner le fait que si la forme fonctionnelle de la partie fixe est linéaire ou approximativement linéaire quelque soit la spécification de l'erreur (v.g. les modèles BC_* et BC_*AU pour ($\lambda_y = \lambda_{xk} = 1$), BC_AAU, BC_AAUHG_1, BC_AAUHG_2 et $BC_{AA}AUHG_{1,2}$), la variable nombre d'automobiles NC semble être la principale contribution à l'explication des déplacements des écoliers, alors que si la forme fonctionnelle est non-linéaire (v.g. les modèles BC_CAUHG_1 et $BC_{CC}AUHG_{1,2}$), la variable temps de déplacement en auto T2T semble posséder le pouvoir explicatif le plus considérable. Par ailleurs, lorsque la forme logarithmique (avec correction pour l'autocorrélation) est utilisée, les résultats semblent suggérer que la variable FSCD est primordiale dans le choix du mode de transport.

Observons maintenant la variation dans la grandeur des élasticités.

Malgré quelques écarts, la variable tarif FSCD semble la plus robuste aux modifications dans les paramètres de la forme fonctionnelle. L'élasticité-

tarif (élasticité propre) varie entre -.15 et -.58 selon les différentes variantes du modèle. Ainsi, dans le modèle où nous libérons la contrainte d'égalité entre les λ_{xk} et où nous corrigeons pour l'hétéroscédasticité (*TEND* est choisi pour X_m), l'élasticité-tarif suggère que le transport en commun est un bien plus inélastique que ne le laissent voir les autres modèles. Ceci signifie qu'une firme qui détiendrait le monopole du marché ou encore qu'un producteur de services qui serait libre d'imposer les prix qu'il désire (absence de réglementation) pourraient accroître leurs prix (jusqu'à ce que l'élasticité propre égale l'unité) et ainsi maximiser leurs profits.

L'élasticité de la demande par rapport au nombre d'automobiles NC (élasticité croisée) est, par ailleurs, assez sensible au comportement de la forme fonctionnelle de la partie fixe et de la spécification de l'erreur. Nous devons mentionner que la plupart des valeurs obtenues pour l'élasticité-nombre d'automobiles NC sont passablement exagérées. Considérons les exemples illustratifs suivants: une augmentation de 10% dans le nombre d'autos par exemple, entraînerait pour les modèles BC_AAUHG_1 ($X_m \equiv$ *SHONF*) et BC_AAUHG_2 ($X_m \equiv$ *TEND*) des baisses respectives de 11% et de 10.4% dans le nombre de passagers du transport en commun; une augmentation similaire, signifierait pour le modèle $BC_{AA}AUHG_{1,2}$ ($X_m \equiv$ *SHONF* et *TEND*) une diminution de 28.4%. Par contre, une hausse de 10% dans le nombre d'automobiles provoquerait une diminution de 13.2% dans le nombre d'usagers pour le modèle BC_CHG_2 et une baisse de 5.5% pour le modèle BC_EHG_2. Conformément, l'élasticité-nombre d'autos pour la plupart des autres modèles montre que la diminution des usagers du transport public est proportionnellement plus élevée que la hausse du nombre d'automobiles. Ces résultats sont très contrariants car ils signifient que le transport en commun ne jouit d'aucune stabilité de la clientèle étudiante face à ce facteur qui tend à le désavantager.

L'élasticité-prix croisée P, a un effet plus ou moins appréciable sur le comportement des enfants ou plutôt sur celui de leurs parents, selon la spécification de la forme fonctionnelle des parties fixe et aléatoire. Nous pouvons souligner cette affirmation à partir de la comparaison des modèles BC_AAUHG_2 ($\hat{\lambda}_y = \hat{\lambda}_{xk} = 1.004$; $\hat{\rho}_\ell$; $\hat{\lambda}_{e2} = 1.07$; $\hat{\delta}_2 = .12$) et BC_AAUHC_2 ($\hat{\lambda}_y = \hat{\lambda}_{xk} = 1.14$; $\hat{\rho}_\ell$; $\lambda_{e2} = 0$; $\delta_2 = 2$). Pour le premier modèle, nous avons une élasticité-prix évaluée à -.17, alors pour le second, elle s'établit à 2.09 (valeur qui est tout à fait irraisonnable). De cette comparaison, il ressort que le transport en commun peut être considéré comme un bien complémentaire lorsque nous introduisons la formulation générale de l'hétéroscédasticité, dans laquelle $X_m \equiv$ *TEND*, et comme un fort substitut aux autres biens lorsque nous utilisons la formulation classique de l'hétéroscédasticité. Il est à noter cependant, que si la variable *SHONF* intervient dans $f(X_t)$ lorsque nous considérons l'hétéroscédasticité générale, le transport en commun est alors considéré comme un bien substitut. De même, la comparaison des modèles BC_AAUHG_1 ($\hat{\lambda}_y = \hat{\lambda}_{xk} = .75$; $\hat{\rho}_\ell$; $\hat{\lambda}_{e1} = -.46$; $\hat{\delta}_1 = .17$) et $BC_{AA}AUHG_{1,2}$ ($\hat{\lambda}_y = \hat{\lambda}_{xk} = .92$; $\hat{\rho}_\ell$; $\hat{\lambda}_{e1} = 5.38$, $\hat{\lambda}_{e2} = .36$; $\hat{\delta}_1 = -.03$, $\hat{\delta}_2 = .49$) montre des valeurs consécutives de .62 et 1.52 pour l'élasticité-prix. Elle devient très élastique dans le second modèle. Cela signifie que l'usage du transport en commun est un substitut très désirable aux autres biens qui voient accroître leur prix.

Après avoir exploré le pouvoir explicatif des variables de prix, nous allons examiner le comportement des variables de revenu et de temps.

Pour la variable revenu, il est intéressant de remarquer que lorsque la valeur estimée du paramètre λ de la partie fixe est approximativement égale à l'unité, les résultats tendent à surestimer le pouvoir explicatif de cette variable. Pour ces modèles, l'élasticité-revenu est quelquefois supérieure à un ($n > 1$), ce qui signifie que le transport en commun constitue un bien supé-

rieur de luxe. Il est toutefois rassurant, pour la plupart des applications, que le transport public soit considéré comme un bien supérieur nécessaire ($n < 1$).

La plupart des variables de niveau de service présentent des élasticités raisonnables quelles que soient les valeurs acquises par les paramètres impliqués dans les vecteurs λ, δ et ρ. Si nous retenons les modèles où les trois variables présentent les signes anticipés, nous pouvons remarquer que l'élasticité-temps d'attente TW est plus élevée que celle du temps de transport TT, ce qui indique que les écoliers sont avant tout sensibles au temps d'attente. Ceci semble raisonnable car une fois engagés dans le système du transport en commun, les individus réagissent d'une manière différente au temps; ils sont alors captifs du mode choisi. Toujours sur la base des modèles considérés, nous pouvons constater qu'une variation d'une minute dans le temps d'attente crée des impacts allant de 1.3 fois à 5.5 fois (v.g. tableau 16) plus élevés qu'une variation identique dans le temps de transport sur la demande de transport public.

TABLEAU 16

Rapport des élasticités du temps d'attente sur le temps de transport selon divers modèles.

Modèles	BC_*AU (λ=1)	BC_AAU	BC_AAUHG_1	BC_AAUHG_2	BC_BAUHG_2	BC_CAUHG_2	$BC_{AA}AUHG_{1,2}$
TW/TT	1.3	2.8	2.0	5.5	3.0	1.5	3.5

Nous devons ajouter que le taux d'élasticité du temps de trajet est la plupart du temps légèrement inférieur au taux d'élasticité propre du tarif. Cette comparaison suggère que les écoliers sont plus sensibles au prix qu'ils ne le sont aux niveaux de service du mode considéré.

Par ailleurs, une comparaison des élasticités des temps de déplacement entre les deux modes s'avère implicite, car il n'est pas clair à priori que les écoliers soient moins sensibles au temps de transport en autobus qu'au temps de trajet en auto. Cependant, l'élasticité de T2T est toujours plus élevée (elle atteint quelquefois des valeurs absurdes: 3.95 pour BC_CAUHG_1, 2.34 pour BC_CAUHG_2, 7.36 pour BC_EHG_2) que celle du temps d'attente TW, ce qui ne confirme pas nos attentes théoriques.

Nous avons également calculé la valeur implicite des prix du temps pour certains de nos modèles. Ainsi, nous pouvons obtenir quelle variation dans le tarif impliquerait la même perte de passagers qu'une variation d'une heure dans chacune des variables de temps.

Pour fins de comparaison, nous observons les résultats présentés dans le tableau 17.

TABLEAU 17

Valeur implicite des prix du temps*: ($/heure).

Modèles	Prix du temps d'attente	Prix du temps de trajet en autobus	Prix du temps de trajet en auto
BC_*AU (λ=1)	.44	.06	.35
BC_AAU	.61	.04	.61
BC_AAUHG_1	.52	.05	.55
BC_AAUHG_2	.74	.02	.72
BC_AAUHC_2	.40	.35	.21
BC_BAUHG_2	1.00	.06	.55
BC_CAUHG_1	2.62	.09	4.06
BC_CAUHG_2	.68	.08	1.76
$BC_{AA}AUHG_{1,2}$	.95	.03	.58

* Notons que les prix du temps sont obtenus par le rapport suivant:

$$\frac{\eta_t}{\eta_p} \cdot \frac{\bar{p}}{\bar{t}} \cdot 60 \quad ,$$

où η_t est le taux d'élasticité propre par rapport au temps (temps d'attente ou temps de trajet en autobus) ou encore le taux d'élasticité croisée de la demande par rapport au temps de transport en automobile; η_p est le taux d'élasticité propre par rapport au prix (tarif); $\bar{p}$ et $\bar{t}$ représentent respectivement les valeurs moyennes du prix et du temps (temps d'attente ou temps de trajet).

Les résultats obtenus indiquent que les écoliers seraient prêts à payer de \$ 0.40 (BC_AAUHC_2) à \$ 2.62 (BC_CAUHG_1) de plus pour attendre une heure de moins en moyenne par déplacement. Par contre, ces mêmes étudiants seraient prêts à payer de \$ 0.02 (BC_AAUHG_2) à \$ 0.35 (BC_AAUHC_2) pour demeurer une heure de moins en autobus. Il est intéressant de remarquer, à travers ces résultats, que l'introduction de la formulation classique de l'hétéroscédasticité semble produire des variations diamétralement opposées à celles des autres modèles.

Par ailleurs, les valeurs implicites du temps de trajet en automobile ne justifient pas toujours les observations faites sur les taux d'élasticités.

La plupart de nos résultats (tableaux 14 et 17) semblent indiquer que si le budget de la Commission de Transport était déficitaire, les autorités devraient opter pour une baisse des fréquences plutôt que pour une hausse des tarifs, les écoliers étant plus sensibles à l'argent qu'au temps. (Cette remarque est basée sur les seuls résultats obtenus des flux de passagers de la clientèle étudiante; des résultats inverses sont anticipés pour la clientèle adulte.)

Les résultats présentés au tableau 14 nous montrent également l'ins-

tabilité des coefficients autorégressifs face aux paramètres de la forme fonctionnelle et des autres paramètres. Ils nous instruisent sur l'importance d'utiliser une spécification complexe de la structure autorégressive et par conséquent, il aurait été injustifié de formuler un modèle en différences premières quelle que soit l'application choisie.

Les variations dans les valeurs numériques obtenues pour les élasticités traduisent encore une fois l'importance de ne pas choisir une formulation arbitraire de la forme fonctionnelle des parties fixe et stochastique mais, d'utiliser pour notre fonction de demande une formulation qui "reflète une plus grande réalité" même si les élasticités obtenues semblent parfois exagérées et non représentatives de la théorie économique.

A cause du peu d'informations(manque d'observations, par exemple) fourni par le modèle, nous ne pouvons pas affirmer qu'une application est préférable à une autre sur la base du critère des élasticités; en fait, le seul critère qui peut nous permettre d'adopter un modèle plutôt qu'un autre est la valeur du logarithme de la fonction de vraisemblance. Néanmoins, nos résultats nous permettent de constater que certaines élasticités semblent plus raisonnables que certaines autres dans l'explication des déplacements des écoliers dans le réseau. Si nous considérons le modèle qui optimise la fonction de vraisemblance, nous pouvons constater que certaines variables (NC et RAWEM entre autres) présentent des élasticités incompatibles avec la théorie même si elles sont statistiquement significatives, alors que d'autres qui ont des valeurs de L^{***} plus faibles présentent des élasticités raisonnables mais non significatives.

CONCLUSION

Les résultats ex post des expériences diverses semblent aboutir à une justification équilibrée en faveur de l'utilité de la détermination simultanée de la forme fonctionnelle des parties systématique et stochastique du modèle.

Il a été démontré que la forme fonctionnelle de l'hétéroscédasticité générale des résidus peut, tout comme la forme fonctionnelle de la partie fixe, refléter une amélioration des modèles à estimer à travers une meilleure spécification.

Nos résultats montrent que la transformation Box-Cox de la partie systématique peut modifier le caractère hétéroscédastique des résidus en agissant comme outil de compensation pour l'hétéroscédasticité des erreurs et que de ce fait, l'hétéroscédasticité aussi bien que la non-linéarité est un élément qui affecte la forme fonctionnelle. La distinction entre les deux éléments qui agissent sur la forme fonctionnelle est rendue possible par l'introduction de la formulation générale de l'hétéroscédasticité. Dans ce dernier cas, seule la non-linéarité influence la forme fonctionnelle; l'hétéroscédasticité y est évanescente.

Sur la base de notre échantillon, nous avons également démontré, à travers les diverses applications, que l'estimation de la forme fonctionnelle optimale juxtaposée à une matrice variance-covariance des erreurs résiduelles qui incorpore à la fois la formulation générale de l'hétéroscédasticité et une structure d'autocorrélation d'un ordre multiple crée des impacts sur la vraisemblance de l'échantillon, sur la signification des variables, sur les signes des paramètres et sur les élasticités dérivées de la demande. Etant donné les variations existantes à ces niveaux, nous voyons donc l'importance de ne pas

déterminer des hypothèses arbitraires sur la forme fonctionnelle des parties fixe et stochastique du modèle.

Nos estimés montrent que les résultats les "meilleurs" sont inhérents au modèle qui optimise la valeur logarithmique de la fonction de vraisemblance. Nous rappelons que les élasticités de certaines variables (nombre d'automobiles, prix et revenu) de ce modèle sont cependant irréalistes.

Lorsqu'un modèle possède un pouvoir prévisionnel, il est primordial d'utiliser une spécification qui génère les "meilleurs" résultats. Les élasticités-prix et niveau de service se doivent d'être aussi justes et raisonnables que possible car elles sont d'un intérêt particulier pour les agents politiques et les preneurs de décisions.

Finalement, nous pouvons ajouter, sur la base du critère de la vraisemblance de l'échantillon, que l'interaction de deux variables peut avoir une influence considérable sur la valeur de la fonction de vraisemblance.

APPENDICE A

LOGARITHME DE LA FONCTION DE VRAISEMBLANCE CONCENTREE.

Nous pouvons réécrire l'expression (35), en lui substituant les termes $[\phi_t]^{\sim}$ et $[\phi_t]^{\tilde{**}}$ tels que définis respectivement dans (34) et (36). Ainsi,

$$\mathcal{L}(y;\ \beta_k,\ \sigma^{**2},\ \rho_\ell,\ \lambda_y,\ \lambda_{xk},\ \lambda_{em},\ \delta_m) = \prod_{t=1+r}^{n} \frac{y_t^{(\lambda_y-1)}}{\sqrt{2\pi}\ \sigma^{**}\ \sqrt{\exp(\Sigma_m \delta_m X_{m_t})^{(\lambda_{em})}}} \cdot$$

$$\exp\left\{ -\frac{1}{2\sigma^{**2}} \left[\left(\frac{y_t^{(\lambda_y)}}{\sqrt{\exp(\Sigma_m \delta_m X_{m_t})^{(\lambda_{em})}}} - \sum_{\ell=1}^{r} \rho_\ell \frac{y_{t-\ell}^{(\lambda_y)}}{\sqrt{\exp(\Sigma_m \delta_m X_{m_{t-\ell}})^{(\lambda_{em})}}} \right) - \Sigma_k \beta_k \right.\right.$$

$$\left.\left. \left(\frac{X_{k_t}^{(\lambda_{xk})}}{\sqrt{\exp(\Sigma_m \delta_m X_{m_t})^{(\lambda_{em})}}} - \sum_{\ell=1}^{r} \rho_\ell \frac{X_{k_{t-\ell}}^{(\lambda_{xk})}}{\sqrt{\exp(\Sigma_m \delta_m X_{m_{t-\ell}})^{(\lambda_{em})}}} \right) \right]^2 \right\} . \qquad \text{(A-1)}$$

a) Le logarithme de la fonction de vraisemblance <u>non concentrée</u> est

$$L(y;\ \beta_k,\ \sigma^{**2},\ \rho_\ell,\ \lambda_y,\ \lambda_{xk},\ \lambda_{em},\ \delta_m) =$$

$$L^* = \sum_{t=1+r}^{n} \left\{ (\lambda_y - 1)\ \ell n\ y_t - \frac{1}{2}\ \ell n\ \sigma^{**2} - \frac{1}{2}\ \ell n\ f(X_t) \right.$$

$$\left. - \frac{1}{2\sigma^{**2}} [y_t^{**} - \Sigma_k \beta_k X_{k_t}^{**}]^2 \right\} , \qquad \text{(A-2)}$$

où y_t^{**} et $X_{k_t}^{**}$ sont définis dans les relations (38) et (40).

Etant donné que $-\frac{1}{2}\ \ell n\ \sigma^{**2}$ peut être considéré comme une constante

pour une somme de 1 + r jusqu'à n, alors (A-2) devient

$$L^* = (\lambda_y - 1) \sum_{t=1+r}^{n} \ln y_t - \left(\frac{n-r}{2}\right) \ln \sigma^{**2} - \frac{1}{2} \sum_{t=1+r}^{n} \ln f(X_t)$$

$$- \frac{1}{2\sigma^{**2}} \sum_{t=1+r}^{n} [y_t^{**} - \Sigma\beta_k X_{k_t}^{**}]^1 \qquad \text{(A-3)}$$

b) Logarithme de la fonction de vraisemblance concentrée sur σ^{**2} et sur β.

. Concentration sur σ^{**2}.

En prenant la dérivée première de (A-3) par rapport à σ^{**2}, nous obtenons:

$$\frac{\partial L^*}{\partial \sigma^{**2}} = - \left(\frac{n-r}{2}\right) \frac{1}{\sigma^{**2}} + \frac{1}{2(\sigma^{**2})^2} \sum_{t=1+r}^{n} [y_t^{**} - \Sigma\beta_k X_{k_t}^{**}]^2 \quad . \qquad \text{(A-4)}$$

En posant cette dérivée égale à zéro, nous avons:

$$\hat{\sigma}^{**2} = \frac{1}{(n-r)} \sum_{t=1+r}^{n} [y_t^{**} - \Sigma\beta_k X_{k_t}^{**}]^2 \quad . \qquad \text{(A-5)}$$

Le logarithme de la fonction de vraisemblance concentrée sur σ^{**2} est obtenu en remplaçant σ^{**2} dans (A-3) par $\hat{\sigma}^{**2}$ tel qu'exprimé dans (A-5); celui-ci est égal à:

$L(y; \beta_k, \rho_\ell, \lambda_y, \lambda_{xk}, \lambda_{em}, \delta_m) =$

$$L^{**} = (\lambda_y - 1) \sum_{t=1+r}^{n} \ln y_t - \left(\frac{n-r}{2}\right) \ln \hat{\sigma}^{**2} - \frac{1}{2} \sum_{t=1+r}^{n} \ln f(X_t)^2 \quad . \qquad \text{(A-6)}$$

1.- Le terme $\sqrt{\pi}$ n'apparaît pas dans le logarithme de la fonction de vraisemblance puisqu'il n'intervient pas dans le processus de maximisation (Tran Cong Liem, 1979, p.2).

2.- La substitution de $\hat{\sigma}^{**2}$ dans (A-5) permet d'abréger cette fonction de sa dernière composante.

. Concentration sur β.

Après avoir concentré la fonction logarithmique sur σ^{**2}, nous devons également la concentrer sur les paramètres β_k. Pour parvenir à une telle concentration, nous devons utiliser la notation matricielle de l'expression (A-5). Par conséquent, nous pouvons écrire que

$$\hat{\sigma}^{**2} = \frac{1}{(n-r)} \sum_{t=1+r}^{n} [w]^2 = \frac{1}{(n-r)} w'w \quad , \qquad (A\text{-}7)$$

où $\underset{((n-r)\times 1)}{w} = \underset{((n-r)\times 1)}{y^{**}} - \underset{((n-r)\times p)}{X^{**}} \underset{(p\times 1)}{\beta}$.

Exprimé en termes de β, (A-7) devient donc:

$$\hat{\sigma}^{**2} = \frac{1}{(n-r)} (y^{**} - X^{**}\beta)' (y^{**} - X^{**}\beta) \quad . \qquad (A\text{-}8)$$

La dérivée première de l'expression (A-6) par rapport à β est

$$\frac{\partial L^{**}}{\partial \beta} = - \left(\frac{n-r}{2}\right) \frac{1}{\hat{\sigma}^{**2}} \cdot \frac{\partial \hat{\sigma}^{**2}}{\partial \beta}$$

$$= - \frac{1}{\hat{\sigma}^{**2}} \left[\frac{\partial w'}{\partial \beta}\right] w$$

A partir de (A-5), nous pouvons démontrer que

$$\sum_{t=1+r}^{n} [y_t^{**} - \Sigma_k \beta_k X_{k_t}^{**}]^2 = (n-r)\ \hat{\sigma}^{**2} \quad .$$

Ainsi, en substituant ce dernier terme dans la relation (A-3), celui-ci se simplifie à $- \frac{1}{2\hat{\sigma}^{**2}} (n-r)\ \hat{\sigma}^{**2} = - \left(\frac{n-r}{2}\right)$

ce qui équivaut à une constante.

$$= - \frac{1}{\hat{\sigma}^{**2}} \left[\frac{\partial (y^{**} - X^{**}\beta)'}{\partial \beta} \right] (y^{**} - X^{**}\beta)$$

$$= - \frac{1}{\hat{\sigma}^{**2}} (- X^{**'}) (y^{**} - X^{**}\beta)$$

$$= \frac{1}{\hat{\sigma}^{**2}} (X^{**'} y^{**} - X^{**'} X^{**}\beta) \quad . \tag{A-9}$$

En égalisant cette expression à zéro, nous obtenons:

$$\hat{\beta}(\rho, \lambda_y, \lambda_x, \lambda_e, \delta) = (X^{**'} X^{**})^{-1} X^{**'} y^{**} \quad , \tag{A-10}$$

lequel est équivalent à l'estimateur des moindres carrés ordinaires.

En remplaçant dans la relation (A-8) β par $\hat{\beta}$, nous obtenons:

$$\hat{\sigma}^{**2}(\rho, \lambda_y, \lambda_x, \lambda_e, \delta) =$$

$$\frac{1}{(n-r)} [y^{**} - X^{**}\hat{\beta}(\rho,\lambda_y,\lambda_x,\lambda_e,\delta)]' [y^{**} - X^{**}\hat{\beta}(\rho,\lambda_y,\lambda_x,\lambda_e,\delta)] \quad . \tag{A-11}$$

Le logarithme de la fonction de vraisemblance concentrée sur σ^{**2} et sur β, est donné par la substitution de (A-11) dans (A-6):

$$L(y; \rho, \lambda_y, \lambda_x, \lambda_e, \delta) =$$

$$L^{***} = (\lambda_y - 1) \sum_{t=1+r}^{n} \ln y_t - \left(\frac{n-r}{2}\right) \ln \hat{\sigma}^{**2}(\rho,\lambda_y,\lambda_x,\lambda_e,\delta)$$

$$- \frac{1}{2} \sum_{t=1+r}^{n} \ln f(X_t) \quad . \tag{A-12}$$

APPENDICE B

HYPOTHESE DE NORMALITE DU TERME STOCHASTIQUE.

Réécrivons l'équation (48) en termes de y_t^{**} et $X_{k_t}^{**}$. Après transformation par les égalités (38) à (41), la relation (48) s'écrit de façon condensée comme étant:

$$y_t^{**} = \Sigma_k \beta_k X_{k_t}^{**} + w_t \quad . \tag{B-1}$$

Pour que la transformation Box-Cox reste valide, en termes de la variable originale y_t, le terme w_t doit respecter une certaine condition.

Afin de dériver la condition essentielle pour w_t, nous devons reformuler l'équation (B-1) autrement. Soit,

$$y_t^{*} = \Sigma_\ell \rho_\ell y_{t-\ell}^{*} + \Sigma_k \beta_k X_{k_t}^{**} + w_t \tag{B-2}$$

$$y_t^{(\lambda_y)} = \sqrt{f(X_t)} \left(\Sigma_\ell \rho_\ell y_{t-\ell}^{*} + \Sigma_k \beta_k X_{k_t}^{**} + w_t\right) \quad . \tag{B-3}$$

En prenant l'inverse de cette transformation, nous obtenons:

$$y_t = \left[1 + \lambda_y \sqrt{f(X_t)} \left(\Sigma_\ell \rho_\ell y_{t-\ell}^{*} + \Sigma_k \beta_k X_{k_t}^{**} + w_t\right)\right]^{(1/\lambda_y)} \quad . \tag{B-4}$$

Comme la transformation monotonique Box-Cox s'applique seulement aux observations réelles positives (donc pour $y_t > 0$), alors le membre de droite de l'expression (B-4) est strictement plus grand que zéro. Conséquemment,

$$\left[1 + \lambda_y \sqrt{f(X_t)} \left(\Sigma_\ell \rho_\ell y_{t-\ell}^{*} + \Sigma_k \beta_k X_{k_t}^{**} + w_t\right)\right] > 0 \quad . \tag{B-5}$$

La condition pour w_t qui nous permet d'obtenir $y_t > 0$ est différente

selon que λ_y est plus grand que zéro ou que λ_y est plus petit que zéro.

. Si $\lambda_y > 0$, il faut que $w_t > w_t^*$.

En effet, si $\lambda_y > 0$ ceci implique que

$$[1 + \lambda_y \sqrt{f(X_t)} \; (\Sigma_\ell \rho_\ell \, y_{t-\ell}^{\,*} + \Sigma_k \beta_k X_{k_t}^{\,**} + w_t)] > 0$$

$$\lambda_y \sqrt{f(X_t)} \; w_t > [-1 - \lambda_y \sqrt{f(X_t)} \; (\Sigma_\ell \rho_\ell \, y_{t-\ell}^{\,*} + \Sigma_k \beta_k X_{k_t}^{\,**})]$$

$$w_t > - \frac{1}{\lambda_y \sqrt{f(X_t)}} - (\Sigma_\ell \rho_\ell \, y_{t-\ell}^{\,*} + \Sigma_k \beta_k X_{k_t}^{\,**}) = w_t^* \qquad \text{(B-6)}$$

$$\underline{w_t > w_t^*} \; . \qquad \text{(B-7)}$$

. Si $\lambda_y < 0$, il faut que $w_t < w_t^*$.

Par une démonstration similaire, il en résulte que

$$w_t < - \frac{1}{\lambda_y \sqrt{f(X_t)}} - (\Sigma_\ell \rho_\ell \, y_{t-\ell}^{\,*} + \Sigma_k \beta_k X_{k_t}^{\,**}) = w_t^* \qquad \text{(B-8)}$$

$$\underline{w_t < w_t^*} \; . \qquad \text{(B-9)}$$

Si les conditions (B-7) et (B-9) sont respectées, alors nous sommes assurés que la variable dépendante y_t telle que définie dans l'expression (B-4) est strictement positive.

En prenant les valeurs anticipées des deux côtés de l'équation (B-1), nous obtenons:

$$E(y_t^{\,**}) = \Sigma_k \hat{\beta}_k X_{k_t}^{\,**} + E(w_t)$$

$$= \hat{y}_t^{\,**} + E(w_t) \; .$$

Le terme résiduel w_t possède une distribution normale "tronquée"

- dans la région $[w_t^*, \infty]$ avec moyenne :

$$E(w_t) = \hat{w}_t^* \int_{-\infty}^{\hat{w}_t^*} f_n\,(0, \hat{\sigma}^{**2})\,dz + \int_{\hat{w}_t^*}^{\infty} z\, f_n\,(0, \hat{\sigma}^{**2})\,dz \qquad \text{(B-10)}$$

si $\lambda_y > 0$,

et

- dans la région $[-\infty, w_t^*]$ avec moyenne :

$$E(w_t) = \int_{-\infty}^{\hat{w}_t^*} z\, f_n\,(0, \hat{\sigma}^{**2})\,dz + \hat{w}_t^* \int_{\hat{w}_t^*}^{\infty} f_n\,(0, \hat{\sigma}^{**2})\,dz \qquad \text{(B-11)}$$

si $\lambda_y < 0$,

où $\hat{w}_t^*$ est la valeur calculée de w_t^* et, $f_n\,(0, \hat{\sigma}^{**2})$ est la fonction de densité normale univariée avec moyenne 0 et variance $\hat{\sigma}^{**2}$.

Par ailleurs, le terme résiduel w_t possède une distribution normale "non-tronquée"

- dans la région $[-\infty, \infty]$ avec moyenne :

$$E(w_t) = 0 \quad , \text{ si } \lambda_y \to 0 \quad . \qquad \text{(B-12)}$$

En effet, si $\lambda_y \to 0$, le membre de droite de la relation (B-4) a la valeur limite de $\exp\,[\sqrt{f(X_t)}\,(\Sigma_\ell\, \rho_\ell\, y_{t-\ell}^{\,*} + \Sigma_k \beta_k X_{k_t}^{\,**} + w_t)]$, laquelle est toujours positive, de sorte que les valeurs observées pour y_t sont également positives pour $-\infty < w_t < \infty$. Donc, puisque le terme d'erreur peut se déplacer à l'intérieur de l'intervalle $-\infty$ et ∞ quand $\lambda_y \to 0$, nous pouvons conclure à l'hypothèse de normalité et ainsi $E(w_t) = 0$. Cette démonstration est exposée dans le texte de Tran Cong Liem (1979, pp.12-14).

APPENDICE C

MESURES DU COEFFICIENT DE CORRELATION MULTIPLE.

. Le premier R^2 relie les erreurs estimées à la valeur non transformée de la variable dépendante. Pour cette raison, il fournit une mesure convaincante de la qualité de l'ajustement obtenue du modèle.

Ce R^2 est exprimé en termes de y_t et de $E(y_t)$:

$$R^2\,(E(y)) = 1 - \sum_{t=1+r}^{n} [y_t - E(y_t)]^2 / \sum_{t=1+r}^{n} (y_t - \bar{y})^2 \qquad (C\text{-}1)$$

où

$$E(y_t) = [1 + \hat{\lambda}_y \sqrt{\hat{f}(X_t)}\,(\sum_{\ell} \hat{\rho}_\ell \hat{y}_{t-\ell}^{\,*} + \sum_{k} \hat{\beta}_k \hat{X}_{k_t}^{\,**} + E(w_t))]^{1/\hat{\lambda}_y} \quad . \qquad (C\text{-}2)$$

L'expression (C-2) provient des valeurs anticipées de l'équation (B-4) de l'annexe B.

Le symbole "$\hat{}$" indique que λ_y, ρ, β et tous les paramètres impliqués dans $f(X_t)$, $y_{t-\ell}^{\,*}$ et $X_{k_t}^{\,**}$ ont été remplacés par leurs estimateurs du maximum de vraisemblance.

. Les deuxième et troisième mesures relient les erreurs estimées à la valeur transformée de la variable dépendante.

Le second R^2 est exprimé en termes de y^{**} et de $E(y^{**})$, alors que le troisième R^2 est exprimé en termes de y^{**} et de $\hat{y}^{**} = X^{**}\hat{\beta}$.

Respectivement, nous avons:

$$R^2\,(E(y^{**})) = 1 - \sum_{t=1+r}^{n} [y_t^{**} - E(y_t^{**})]^2 / \sum_{t=1+r}^{n} (y_t^{**} - \bar{y}^{**})^2 \qquad (C\text{-}3)$$

et

$$R^2(\hat{y}^{**}) = 1 - \sum_{t=1+r}^{n} (y_t^{**} - \hat{y}_t^{**})^2 / \sum_{t=1+r}^{n} (y_t^{**} - \bar{y}^{**})^2 \quad . \text{(C-4)}$$

Il est à noter que dans le cas linéaire sans autocorrélation et sans hétéroscédasticité, les trois mesures sont équivalentes.

Ces mesures sont développées dans le texte de Tran Cong Liem (pp.12-14).

APPENDICE D

DEFINITION DES VARIABLES.

A - VARIABLE DEPENDANTE

XCSCNFA : nombre de voyages effectués par les écoliers.

B - VARIABLES DE PRIX

1. FSCD : tarif réel exigé pour les écoliers âgés de 18 ans et moins par les autorités de la C.T.C.U.M..
2. NC : nombre total d'automobiles dans le territoire desservi par la C.T.C.U.M..
3. P : indice des prix à la consommation dans la région métropolitaine de Montréal.

C - VARIABLE DE REVENU

4. RAWEM : salaire hebdomadaire réel moyen à Montréal.

D - VARIABLES DE TEMPS

5. TW : temps moyen d'attente en un point aléatoire du réseau pendant une journée moyenne d'un mois donné.
6. TT : temps moyen d'un déplacement représentatif par transport en commun sur une distance de quatre milles.
7. T2T : temps moyen d'un déplacement représentatif hors pointe par automobile sur une distance de quatre milles.

E - VARIABLES DE CONFORT

8. TLK : valeur absolue de la différence entre la température mensuelle moyenne subie et celle d'un mois typique particulier, septembre (56 degrés Fahrenheit).
9. RFK : précipitation mensuelle de pluie.
10. SFK : précipitation mensuelle de neige.

11. CSFMR : accumulation mensuelle de neige au sol.

F - VARIABLES D'ACTIVITE

12. RSCMP : inscriptions scolaires aux niveaux primaire et secondaire dans les institutions de la C.E.C.M..

13. SHONF : indice réel des ventes au détail dans le territoire desservi par la C.T.C.U.M..

14. EX7 : nombre de passagers qui sont entrés à la station de métro Ile Ste-Hélène lors de l'Exposition Internationale de 1967.

15. EX8 : nombre de passagers qui sont entrés sur le site de l'Exposition par cette station en 1968.

16. STR : nombre de jours d'arrêt total de service à cause des grèves.

17. TEND : variable artificielle qui sert à simuler la tendance de la population d'âge scolaire. Il s'agit d'un trend linéaire qui atteint un maximum en octobre 1969 et qui diminue après cette période.

G - VARIABLES COMPTABLES, D'AGREGATION ET DE DECISION ADMINISTRATIVE

18. ANT : variable dichotomique qui mesure l'impact du changement tarifaire de mars 1965.

19. ANTL1 : ANT retardée d'une période.

20. ZSC : variable dichotomique qui tient compte du double comptage d'une certaine fraction de la clientèle des enfants avant la suppression des zones externes.

21. WD : nombre mensuel de jours ouvrables.

22. SAT : nombre mensuel de samedis.

23. SHD : nombre mensuel de dimanches et de jours fériés.

24. CHEAT2 : variable dichotomique pour tenir compte du contrôle informatique des renseignements qui concernent la carte d'écolier.

25. STRUCT1 : variable dichotomique qui mesure l'impact des changements du réseau de surface en fonction de son intégration avec le métro en 1967.

26. CARDEX : variable dichotomique pour tenir compte de l'expiration (depuis

UNIVERSITÉ DE MONTRÉAL

1963) des cartes à prix réduits pour certains écoliers pendant la période estivale.

27. C : constante.

BIBLIOGRAPHIE

1.- Anscombe, F.J. and Tukey, J.W., The Criticism of Transformations. Texte présenté devant the American Statistical Association and Biometric Society, 1954, Montreal.

2.- Beach, C.M. and MacKinnon, J.G., "A Maximum Likelihood Procedure for Regression with Autocorrelated Errors", Econometrica, Vol. 46, 1978, nu. 1, pp.51-58.

3.- Box, G.E.P. and Cox, D.R., "An Analysis of Transformations", Journal of the Royal Statistical Society Series B, Vol. 26, 1964, nu. 2, pp. 211-243.

4.- Box, G.E.P. and Jenkins, G.M., Time Series Analysis Forecasting and Control, Holden-Day, San Francisco, 1970.

5.- Box, G.E.P. and Tildwell, P.W., "Transformation of the Independent Variables", Technometrics, Vol. 4, 1962, nu. 4, pp.531-550.

6.- Dagum, C. et Dagum, E.B., Construction de modèles et analyse économétrique. Cahiers de l'Institut de sciences mathématiques et économiques appliquées 8, 1974, nos. 11-12.

7.- Fletcher, R. and Powell, M.J.D., "A Rapidly Convergent Descent Method for Minimization", The Computer Journal, Vol. 6, 1963, pp.163-168.

8.- Gaudry, M.J.I., The Demand for Public Transit in Montreal and Its Implications for Transportation Planning and Cost-Benefit-Analysis. Ph.D. Dissertation, Economics Department, Princeton University, 1973.

9.- Gaudry, M.J.I., An Aggregate Time-Series Analysis of Urban Transit Demand: The Montreal Case, Publication # 6, Centre de recherche sur les transports, Université de Montréal, août 1974, révisé en avril 1975.

10.- Gaudry, M.J.I., A Study of Aggregate Bi-Modal Urban Travel Supply, Demand and Network Behavior Using Simultaneous Equations with Autoregressive Residuals, Publication # 88, Centre de recherche sur les transports, Université de Montréal, novembre 1977, révisé en août 1978.

11.- Gaudry, M.J.I. and Dagenais, M.G., The Use of Box-Cox Transformations in Regression Models with Heteroskedastic Autoregressive Residuals, Publication # 106, Centre de recherche sur les transports, Université de Montréal, août 1978, révisé en avril 1979.

12.- Gaudry, M.J.I. and Wills, M.J., Estimating the Functional Form of Travel Demand Models, Publication # 63, Centre de recherche sur les transports, Université de Montréal, juin 1977, révisé en janvier 1978.

13.- Glejser, H., "A New Test for Heteroskedasticity", Journal of the American Statistical Association, Vol. 64, 1969, nu. 325, pp.316-323.

14.- Goldfeld, S.M. and Quandt, R.E., "Some Tests for Homoskedasticity", Jour-

nal of the American Statistical Association, Vol. 60, 1965, nu. 310, pp.539-547.

15.- Goldfeld, S.M. and Quandt, R.E., Nonlinear Methods in Econometrics, ed. D.W. Jorgenson and J. Walbroeck, Amsterdam, London, North-Holland, 1972.

16.- Heckman, J. and Polachek, S., "Empirical Evidence on the Functional Form of the Earnings-Schooling Relationship", Journal of the American Statistical Association, Vol.69, 1974, nu. 346, pp.350-354.

17.- Johnston, J., Econometric Methods, McGraw-Hill, 2nd. ed., New-York, 1972, (c) 1963, 437 p..

18.- Kadiyala, K.R., "A Transformation Used to Circumvent the Problem of Autocorrelation", Econometrica, Vol. 36, 1968, nu. 1, pp.93-96.

19.- Kau, J.B. and Sirmans, C.F., The Functional Form of the Gravity Model: A New Technique with Empirical Results. Working paper. Department of Finance, University of Illinois. Presented at the North American Meeting of the Regional Science Association in Toronto, 1976.

20.- Kendall, M.G. and Stuart, A., The Advanced Theory of Statistics, Vol. II, Griffin, London, 1967.

21.- Major, F., L'hétéroscédasticité dans les modèles agrégés de demande de transport en commun en séries chronologiques, Publication # 102, Centre de recherche sur les transports, Université de Montréal, juin 1978.

22.- Park, R.E., "Estimation with Heteroskedastic Error Terms", Econometrica, Vol. 34, 1966, nu. 4, p.888.

23.- Ramsey, J.B. and Zarembka, P., "Specification Error Tests and Alternative Functional Forms of the Aggregate Production Function", Journal of the American Statistical Association, Vol. 66, pp.471-477.

24.- Savin, N.E. and White, K.J., "Estimation and Testing for Functional Form and Autocorrelation: A Simultaneous Approach", Journal of Econometrics, Vol. 8, 1978, pp.1-12.

25.- Schlesselmann, J., "Power Families: A Note on the Box and Cox Transformation", Journal of the Royal Statistical Society Series B, Vol. 33, 1971, nu. 2, pp.307-311.

26.- Spitzer, J.J., "The Demand for Money, the Liquidity Trap, and Functional Forms", International Economic Review, Vol. 17, 1976, pp.220-227.

27.- Spitzer, J.J., "A Simultaneous Equations System of Money Demand and Supply Using Generalized Functional Forms", Journal of Econometrics, Vol.5, 1977, pp.117-128.

28.- Theil, H., "Estimates and their Sampling Variance of Parameters of Certain Heteroskedastic Distributions", dans Revue de l'Institut International de Statistique, Vol. 19, 1951, pp.141-147.

29.- Tran Cong, Liem, DSDM: L-1, A Program for Box-Cox Transformations in Regression Models with Heteroskedastic and Autoregressive Residuals, Publication # 134, Centre de recherche sur les transports, Université de Montréal, juin 1979.

30.- Tukey, J.W., "On the Comparative Anatomy of Transformations", Annals of Mathematical Statistics, Vol. 28, 1957, pp.602-632.

31.- Turner, M.E. et al., "Generalized Asymptotic Regression and Nonlinear Path Analysis", Biometrics, Vol. 17, 1961, pp.120-143.

32.- Welland, J.D., Cognitive Abilities, Schooling and Earnings: The Question of Functional Form. Working paper 76-14. Department of Economics, McMaster University, Hamilton, 1976.

33.- Zarembka, P., "Functional Form in the Demand for Money", Journal of the American Statistical Association, Vol. 63, 1968, nu. 322, pp.502-511.

34.- Zarembka, P., Transformation of Variables in Econometrics, ch. 3 of Frontiers in Econometrics, ed. P. Zarembka, 1974, pp.81-104.

UNIVERSITÉ DE MONTRÉAL

REMERCIEMENTS

Je désire exprimer toute ma reconnaissance envers le Professeur Marc Gaudry d'avoir dirigé d'une façon soutenue et éclairée ce rapport de recherche.

Mes remerciements s'adressent également au Professeur Marcel Dagenais pour ses nombreux conseils et pour avoir participé à la correction de ce rapport.

Je tiens également à remercier M. Tran Cong Liem, assistant de recherche, pour son assistance technique.